IJCSIS Vol. 11 No. 2, February 2013
ISSN 1947-5500

International Journal of Computer Science & Information Security


Cornell University
Library


Cogprints


Google scholar


.docstoc
find and share professional documents


ScientificCommons

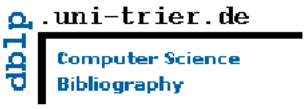

View my documents on
Scribd


BASE
Bielefeld Academic Search Engine


SCIRUS
search engine for science


SciRate.com


CiteSeerx beta


dblp .uni-trier.de
Computer Science
Bibliography


Q·Sensei


DOAJ
DIRECTORY OF
OPEN ACCESS
JOURNALS


EBSCO
HOST


ProQuest

IJCSIS

ISSN (online): 1947-5500

Please consider to contribute to and/or forward to the appropriate groups the following opportunity to submit and publish original scientific results.

CALL FOR PAPERS
International Journal of Computer Science and Information Security (IJCSIS)
January-December 2013 Issues

The topics suggested by this issue can be discussed in term of concepts, surveys, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas.
See authors guide for manuscript preparation and submission guidelines.

Indexed by Google Scholar, DBLP, CiteSeerX, Directory for Open Access Journal (DOAJ), Bielefeld Academic Search Engine (BASE), SCIRUS, Cornell University Library, ScientificCommons, EBSCO, ProQuest and more.

Deadline: see web site
Notification: see web site
Revision: see web site
Publication: see web site

Context-aware systems
Networking technologies
Security in network, systems, and applications
Evolutionary computation
Industrial systems
Evolutionary computation
Autonomic and autonomous systems
Bio-technologies
Knowledge data systems
Mobile and distance education
Intelligent techniques, logics and systems
Knowledge processing
Information technologies
Internet and web technologies
Digital information processing
Cognitive science and knowledge
Agent-based systems
Mobility and multimedia systems
Systems performance
Networking and telecommunications
Software development and deployment
Knowledge virtualization
Systems and networks on the chip
Knowledge for global defense
Information Systems [IS]
IPv6 Today - Technology and deployment
Modeling
Software Engineering
Optimization
Complexity
Natural Language Processing
Speech Synthesis
Data Mining

For more topics, please see web site https://sites.google.com/site/ijcsis/

For more information, please visit the journal website (https://sites.google.com/site/ijcsis/)

Editorial
Message from Managing Editor

*Since 2009, the **International Journal of Computer Science and Information Security** (IJCSIS), has been promoting the dissemination of new knowledge in research areas of computer science and applications, and advances in information security. The themes focus mainly on innovative developments, research issues/solutions in computer science and related technologies. The journal aims at providing a platform and encourages emerging scholars and academicians globally to share their professional and academic knowledge in the fields of computer science*

IJCSIS archives all publications in major academic/scientific databases; abstracting/indexing, editorial board and other important information are available online on homepage. IJCSIS editorial board consisting of international experts solicits your contribution to the journal with your research papers, projects, surveying works and industrial experiences. IJCSIS appreciates all the insights and advice from authors and reviewers. Indexed by the following International agencies and institutions: Google Scholar, Bielefeld Academic Search Engine (BASE), CiteSeerX, SCIRUS, Cornell's University Library EI, Scopus, DBLP, DOI, ProQuest, EBSCO. Google Scholar reported a large amount of cited papers published in IJCSIS.

IJCSIS is currently accepting quality manuscripts for upcoming issues based on original qualitative or quantitative research, an innovative conceptual framework, or a substantial literature review that opens new areas of inquiry and investigation in Computer science. Case studies and works of literary analysis are also welcome.

*We look forward to your collaboration. For further questions please do not hesitate to contact us at **ijcsiseditor@gmail.com**.*

A complete list of journals can be found at:
http://sites.google.com/site/ijcsis/

IJCSIS Vol. 11, No. 2, February 2013 Edition

ISSN 1947-5500 © IJCSIS, USA.

Journal Indexed by (among others):

IJCSIS EDITORIAL BOARD

TABLE OF CONTENTS

infrastructure and processes. These all new policies and practices can have a positive impact on the security of the entire network not only voice communications. This paper provide starting point for understanding the security facets of VoIP in a rapidly evolving set of technologies that are seeing growing deployment and use. The main goal is to provide a better understanding of the security background with respect to VoIP security facet toward directing future research and in other similar up-and-coming technologies.

Keywords— VoIP, ITU-T H.323, Session Initiation Protocol, Media Gateway Control Protocol, Security attacks.

4. Paper 31011312: An Approach To QoS-Aware Web Service Composition Using Learning Automata (pp.25-29)

Ali Mehrpour, Engineering Department, Islamic Azad University, Research and Science Branch, Tehran, Iran
Mir Ali Seyyedi, Engineering Department, Islamic Azad University, Research and Science Branch, Tehran, Iran
Shahrbanoo Majlesi, Engineering Department, Islamic Azad University, Research and Science Branch, Tehran, Iran

Abstract — Because of growing number of alternative web services that provide same functionality with different qualities, how to select and composite web services to satisfy user's end-to-end constraints is a decision problem. In this paper we have proposed an approach for web service composition based on quality parameters using learning automata consists of two steps: Step1) Stochastic Learning Automata for local selection and Step2) Distributed Learning Automata for global optimization to create composite web service. We have applied these to kind of Learning Automata as a part of Broker in Web Service Architecture. Experimental evaluations show our approach can be applied in dynamic web environment with an acceptable performance without any limitation on number of QoS parameters.

Keywords-component; Quality of Service (QoS); Web Service Composition (WSC); Stochastic Learning Automata (SLA); Distributed Learning Automata (DLA); Web Service Architecture

5. Paper 31011317: Demonstration of the Functioning of TCP Protocol Used for Network Congestion Control (pp. 30-35)

Asagba Prince Oghenekaro (1); Anucha Udo Sylvester (1); Ogini Nicholas Oluwole (2)
(1)Faculty of Science, Department of Computer Science, University of Port Harcourt, Port Harcourt, PMB 5323, Choba, Rivers State, Nigeria
(2)Faculty of Science, Department of Mathematics and Computer Science, Delta State University, Abraka, Delta State, Nigeria

Abstract — Congestion can occur when the quality of service in a network reduces as a result of a node or link conveying too many data. TCP is the most widely used protocol for Internet traffic, including email, web browsing, data and an increasing portion of multimedia content delivered in real time using the HTTP/TCP protocols. Performances of existing TCP congestion control algorithms degrade significantly when deployed over wireless networks. TCP was designed primarily for reliability as opposed to real time delivery, but the problem is particularly severe for real time applications, such as, HTTP/TCP based streaming. In this paper, we carried out a research on the TCP's four related congestion control algorithms, namely: slow-start, congestion avoidance, fast retransmit and fast recovery. We studied the behaviour and implementation of slow-start and congestion avoidance algorithms, as well as the modifications to the fast retransmit and fast recovery. We used the OPNET Network Model as our methodology. The TCP performance on the network was modeled, first without background traffic and then with background traffic. We compared these algorithms using the same network model to deterministically check several scenarios; and simulations were conducted to ascertain the differences between the congestion control algorithms studied and OPNET's software. The results gotten showed that using different algorithms, traffic could actually be fine tuned in the network being modeled so as to achieve higher Performance. The adjustments were done in the OPNET simulator.

Keywords - TCP Protocols; Congestion control algorithms;Network; Acknowledgment (ACK); OPNET Network

6. Paper 31011318: Change Management Strategies and Processes for the successful ERP System Implementation: A Proposed Model (pp. 36-41)

Abdullah Saad AL-Malaise AL-Ghamdi
Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University

Abstract— Recent advancement in information technology and business development, the business organizations turned towards the adoption of advanced information technology systems for their organizational setup. Progression of technologies in business environment has been observed in many organizations by the initiation of enterprise resource planning (ERP) system implementation. ERP is business integrated information system software that attracts the attention of business organizations in order to improve their business processes and achieve the company's goals. Almost all the ERP system implementation is based on change management system, where the traditional/ legacy system is completely replaced with the new and advance system. This paper will discuss the change management strategies and processes for the success of ERP system implementation. The paper has proposed a model, change management strategies and processes for the successful ERP system implementation that will strengthen the scope of the title of this paper.

Keywords-component; Change Management, IT, ERP, User Reaction, System, Implementation Process

7. Paper 21011302: Securing AODV with Authentication Mechanism using Cryptographic Pair of Keys (pp. 42-45)

K. Suresh Babu, K. Chandra Sekharaiah)
School of IT, JNT University Hyderabad, India

Abstract -- Mobile Ad Hoc Networks (MANETs) is characterized by self–organizing capability, dynamically configurable infrastructure and multihops. Of late, MANETs form emerging state-of-the-art networking technology faster. The routing protocol plays an important role in it overall operation of MANETs. AODV is one of MANET routing protocol. In this paper, the vulnerabilities in MANETs and security flaws in AODV are discussed. A new security mechanism for securing AODV with message digest authentication using a pair of keys (public key cryptography) is proposed and implemented in NS - 2 simulator.

Keywords – Self-organizing; multihops; authentication; public key

8. Paper 27011306: An Overview of Wireless Local Area Networks (WLAN) (pp. 46-53)

Ibrahim Al Shourbaji, Computer Networks Department, Jazan University, Jazan 82822-6649, Saudi Arabia

Abstract - Wireless Communication is an application of science and technology that has come to be vital for modern existence. From the early radio and telephone to current devices such as mobile phones and laptops, accessing the global network has become the most essential and indispensable part of our lifestyle. Wireless communication is an ever developing field, and the future holds many possibilities in this area. One expectation for the future in this field is that, the devices can be developed to support communication with higher data rates and more security. Research in this area suggests that a dominant means of supporting such communication capabilities will be through the use of Wireless LANs. As the deployment of Wireless LAN increases well around the globe, it is increasingly important for us to understand different technologies and to select the most appropriate one . This paper provides a detailed study of the available wireless LAN technologies and the concerned issues ,will give a brief description of what wireless LANs are ,the need of Wireless LAN ,History of wireless LAN , advantages of Wireless Networks ,with summarizing the related work on WLAN in academic area , Wireless LAN technologies , some risks attacks against wireless technologies , suggesting some recommendations to protect wireless LAN network from attack , Finally we propose some research issues should be focused on in the future.

Keywords: *Wireless Networking, Security, 802.11 Standard, Network security,*

9. Paper 31011319: Ambient Noise Coherence Properties Detection for various Hydrophone Spacing (pp. 54-58)

V.G.Sivaumar, Department of ECE, Sathyabama University, Chennai, India
Dr.V.Rajendran, Department of Physics/Ece, SSN College of Engineering, Chennai, India

Abstract — Ambient noise is a complex and important phenomenon which greatly affects the listening capacity of instruments in underwater environment. The ambient noise in sea is the overall combination of wind speed, wave speed, wave height, barometric pressure, dew point, temperature, marine life, shipping traffic and seismic activities. The present work concentrates on coherence with various hydrophone spacing. Under water ambient noise analysis is essential to enhance the Signal to Noise Ratio (SNR) of acoustic based underwater instruments. This paper investigates the effect of noise spectrum over a different hydrophone spacing and the signal coherence with hydrophone spacing is examined in the Bay of Bengal Sea region.

Keywords-component; Ambient noise; Noise Level; Wind speed; Coherence.

10. Paper 31011313: Adaptive Iris Localization and Recognition: Modification On Daugman's Algorithm (pp. 59-71)

Marwan AL-abed Abu-zanona, Department of Computer Science, Imam Muhammad Ibn Saud Islamic University, KSA
Bassam M. El-Zaghmouri, Department of Computer Information Systems, Jerash University, Jordan

Abstract — The use of biometric information has been widely known for both people identification and security application. It is common knowledge that each person can be identified by the unique characteristics of one or more of biometric features. One most unique and identifiable biometric characteristics is the iris, wherever the second is the voice, and the third is finger print. This research attempts to apply iris recognition techniques based on the technology invented by Dr. John G. Daugman, an attempt of implementing a build an end user application. Iris Recognition is expected to play a major role in a wide range of applications in which a person's identity must be established or confirmed in high reliability and high privacy, Including access controls, authorizations, ID detection, etc. This research depends on standard iris images was token from CASIA database. The most efficient computer language for simulation and technical computing (MATLAB) will be used to make the problem statement and result in addition to mathematical and AI modelling more easier and reliable.

Keywords— Image Processing; Iris; Localization; Biometrics; Gradient

11. Paper 31071240: Design and Implementation of Security Framework for Cognitive Radio Networks Resource Management (pp. 72-86)

Obeten O. Ekabua & Ifeoma U. Ohaeri
Department of Computer Science, North-West University, Mafikeng Campus, Private Bag X2046, Mmabatho 2735, South Africa

Abstract --- Designing and implementing a secure communication for any network is an important issue for the optimal control of resource usage in a resource constrain network environment. Therefore, in this paper, we design and implement a joint authentication and authorization framework by transforming the framework requirement analysis. The framework is a security infrastructure capable of monitoring and controlling access to the limited spectrum resources, dynamically managing data and information in CRN, for a secured communication and quality of service (QOS). We explained how the various components in the framework interact to ensure a secured communication and effective access control.

Keywords--Network Management; security; authentication; authorization; access control.

Smartphones Resources Analysis During Playback of Progressive Video over Wi-Fi

Ali H. Mohammed
Computer Science Dept.
College of Computer Science and mathematics
Iraq, Mosul, Mosul University

Dr. Omar A. Ibrahim
Computer Science Dept.
College of Computer Science and mathematics
Iraq, Mosul, Mosul University

***Abstract*— This paper presents the implementation of progressive video stream to mobile phone over Wi-Fi channel with different CODECs . J2ME is the language that will be adopted, especially the techniques named (MMAPI) which specialized in multimedia technologies in mobile phones. Moreover, the paper will make an analysis of the CPU and RAM resources due to the effect of these resources when playing progressive video stream in a mobile device. The choice of these two resources is made because they directly affect the mobile performance when dealing with different services, especially when using multitasking extensively. Also the paper will make a power consumption analysis in mobile phone when utilising progressive streaming service due to the fact that mobile phone derives energy from a limited lifetime battery depending on its size and mobile activity. The main components of the architecture are HTTP server, Wi-Fi infrastructure, mobile client-enabled Java working under Symbian OS.**

Keywords: *Progressive video stream, Mobile phone, Wi-Fi, CODEC, J2ME, MMAPI, CPU&RAM , Power consumption , HTTP server, Symbian OS*

I. INTRODUCTION

As many people know what the mobile phones are, they do not always realize how to differentiate them from smartphones. In simple words, smartphones are mobile phones having an operating system and equipped with development features such as (WLAN, hard disk, etc.) It is worth mentioning that the first smartphone designed by IBM was named SIMON.

Mobile communications systems have been developed rapidly so that we can seek this evolution day after day and certainly correspond this development with the emergence of new services and applications designed to serve the users. Among these services is streaming multimedia, mainly progressive streaming[1].

Transfering the video/audio file over the network can be done in two methods, downloading and streaming. The size of storage device as well as the available bandwidth play an important role in the transportation process, especially if the file size is fairly large. Downloading needs a time range from minutes to hours while streaming reduces this time to a few seconds for both buffering and playback multimedia[2]. Streaming is an important and interesting service. It can be defined as the transmission of video images from one location on network called video server to another side called client without transfering a single video file. Thus, the video frames are consumed in the client side while the downloading process in progress and eventually the video frames can be viewed to the user as it arrives before all video has been transmitted[3]. This technique (Streaming) offers great facilities for mobile users that are limited in resource (i.e. processor speed, storage, battery. etc.). To illustrate this facility consider the following scenario: Imagine that there is a video file in the network with a size of 100MB, and in order to watch this file, the client needs to load the complete file, which may take a long time for loading (depending on the bandwidth used) in addition to exhaust mobile phone resources that differ from phone to another (processor, memory, storage) and eventually consume power which if it runs out, the mobile phone cannot continue to operate. The client can watch a 100MB video file just after several seconds via Streaming technique [4].

Recently, many wireless technologies have begun to appear. These technologies provided facilities for users to connect their computing device with a wide spectrum of devices in an easy and flexible manner. WLAN, especially Wi-Fi, has appeared as a much more powerful and flexible alternative than wired LAN. However, nowadays mobile manufactures equip their products with these new technologies as an additional connectivity tool.

J2ME or (Java 2 Micro Edition) is a version of the sun micro system's. "J2ME isn't a specific piece of software or specification, all it means is Java for small devices. Small devices range in size from pagers, mobile phones, and personal digital assistants (PDAs)"[5]. J2ME is a part of java 2 which makes with java SE and a java EE, java family that works under JCP(Java Community Process). J2ME appeared in Java One developer Conference in 1999 and the main architecture of this language is represented by three components: Configurations , Profiles and Optional packages[6].

Multimedia on mobile phone running java is handled by a special library called Mobile Media Application Programing Interface (MMAPI) of Java specific request JSR135. It provides a simple and flexible framework for playback audio and video through two steps: the first is ***Protocol handling*** which is concerned with retrieving the media content from a source such as local storage, database, or streaming server and feeds the content to the media-handing system, and the second is Media content handling that parses, decodes and renders the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

media content to the output subsystem such as the audio speaker and display screen[7].

There are two kinds of streaming media: live and progressive. In the first case the client downloads video frames with a speed very close to the bitrate of the source video, so the video frames are received, decoded and displayed in a real time fashion. Live streaming requires a significant amount of computing resources which is limited in mobile phones and is often a specific hardware support. This concept is used in a standard television broadcasting. In contrast to progressive streaming, the video file already exists(stored file), and the users download the file with the highest potential speed between server and client, depending on the server sending capacity and the available bandwidth. In this way the client can play out the video while parts of the video are being received and decoded. The video files are stored at the server and delivered to one or multiple clients when requested (on-demand). Thousands of sites provide streaming of stored audio and video today, including Microsoft Video, YouTube, Vimeo and CNN [8-11].

The proposed work uses the second type of streaming above which make use of on-demand concept . On-demand streaming is activated by the user request and can be presented at any time in accordance with requests from the client, and the user can seek the position of the playback as he/she wishes during watching[3]. This chosen is made because progressive streaming handled by the (HTTP) protocol which is considered to be mandatory included in all mobile phone, in contrast to (Live Streaming)which is originally used with the (RTSP) protocol which is considered to be optional included in mobile phones . Several challenges are to be faced such as the terminal mobile phones heterogeneity, since they have a wide range of capabilities such as the CPU speed, memory size and display resolution. As well as the wireless limitation represented by noise and converge area.

To accomplish the work, the researcher uses Hyper Text Transfer Protocol (HTTP) to receive the video stream from HTTP server as well as different compression techniques to overcome the bandwidth limitation. Also, the work adopts a framework of mobile phone specification to cope with the terminals heterogeneity.

This paper presents the implementation of progressive video stream to mobile phone over Wi-Fi channel and different video CODECs(H.263,H.264,MPEG-4).The audio CODEC used is (MP3). Moreover, this work provides a mobile phone resource analysis during receiving and playing back progressive video stream in mobile phone under the configurations above. The resources under investigation are CPU , RAM, and Battery consumption .The rest of this paper is organized as follows: Section two presents related work. Section three explains the main architecture of the proposed system. Section four shows the test performed, the measurements and the results. Finally, section five provides the concluding remarks and future research directions.

II. RELATED WORK

Many previous works that are concerned with streaming service used RTSP to achieve streaming, but this protocol is not included in all mobile phones.

The work proposed by Ary Mazharuddin, Shiddiqi Henry and et.al[4], presented an application called "POCKET VIDSTREAM " which is used for playing video stream on mobile phone. The researchers use in this application an on-demand streaming concept which makes use of (RTSP & HTTP server).

Another previous work done by Wang Zhong-rong and Liu Zhao[2] proposes a mobile streaming system that is based on four components: server, client, channel and content source. H.264 has been used as video CODEC and QCP as audio CODEC where the transmission channel is CDMAx.

Mabel Vazquez-Briseno and Pierre Vincent[1]presented an adaptable architecture for mobile streaming application. They describe the challenges that face the designer in developing mobile application. 3GPP network has been used since it has a standardized streaming service and specifies both protocol and CODEC. They adopt MPEG-4 as video CODEC and AMR as audio CODEC.

Finally , the work by Eklof[12], aimed to detect the available bandwidth on the client side that is connected to the server via cellular network. Depending on the available network, the streaming server increases and adjusts the video quality using RTP(Real Time Protocol) as the transmitting protocol.

III. SYSTEM COMPONENTS

The main goal of the proposed work is to make it easier for the client to use mobile phone for receiving and displaying progressive video stream over Wi-Fi channel and to provid a complete analysis of the mobile device resource during receiving and playing back video stream . To complete this task, the researchers needed to develop a generic architecture that fits a variety of mobile devices. Figure (1) describes the main component of the proposed system. In order to describe

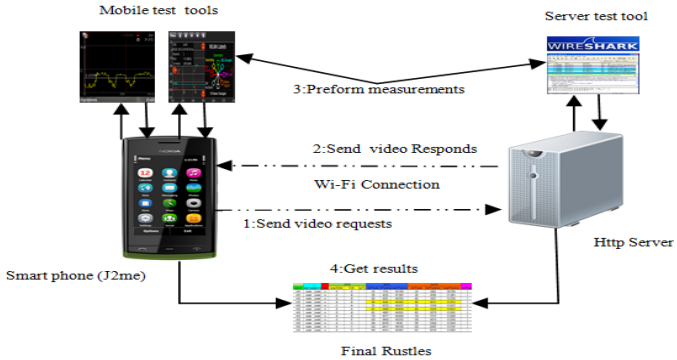


Figure 1: System Architecture

the basic operation that is done in the architecture, we can barely say that the mobile phone (client side) will be connected to the HTTP server by sending a request video command to get progressive video stream. The HTTP server responds with video packets to mobile phone if there is no error in connection.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A. HTTP SERVER

The Hypertext Transfer Protocol (HTTP) is an application-level protocol for distributed, collaborative, hypermedia information systems. HTTP has been in use by the World-Wide Web global information initiative since 1990. In the case of progressive streaming (HTTP streaming) the media file is downloaded as ordinary web pages, but the play out begins just as soon as the first bytes are received (excluding client side buffering) instead of waiting for the entire file to be downloaded. This approach is widely used by video sharing sites on the Internet, such as YouTube[13, 14].

The HTTP server used in this paper is the APACHI server. The Apache HTTP Server Project is a collaborative software development effort aimed at creating a robust, commercial-grade, featureful , and freely-available source code implementation of an HTTP (Web) server[15]. According to the Netcraft survey, APACHI is the most widely used server, where the percentage of deploying this server across the world from December 2011 to January 2012 is around (64.91%) [16]

B. CODEC TECHNOLOGY

One more duty related to the server, as its known that in order to send multimedia data, that needs high bandwidth channel in small bandwidth channel, special technique is required which is known as CODEC. CODEC stands for the compression and decompression used to reduce the amount of redundancy data sent over network. Three types of CODEC adopted in our proposed architecture depends on mobile phone support:

- **MPEG-4:** The Moving Picture Experts Group (MPEG) is an ISO / IEC working group, which was established to define the standards for digital video and audio formats. MPEG-4 was developed to enable the encoding of the rich multimedia content, extending beyond video and audio and also includes vector graphics and similar content. Data rates supported by MPEG-4 range from 10 kbps to 1,000,000 kbps, which makes it ideal for almost any type of video application[17].
- **H.263 :**ITU-T H.263 is an established codec used in various multimedia services. Almost all mobile phones support this type of CODEC and for this reason, the H.263 Profile 0, Level 10 (also known as "H.263 baseline"), has been defined as a mandatory CODEC in mobile devices. It is also a main stream CODEC supported by Nokia video players. H.263 uses the Discrete Cosine Transform (DCT) to reduce spatial redundancy [17] .
- **H.264:** H.264/AVC is the newest international video coding standard. The main goals of this coding technique are to develop a simple and straightforward video coding design, that enhanced compression performance, and to provide a "network-friendly" video representation[18].

C. SMARTPHONE (J2ME Client)

Java language started with one version known as java 2 slandered edition (J2SE). After two years of introducing J2SE in 1995, a new version of java was released namely Java 2 enterprise edition (J2EE).The most recent edition of java family is called Java 2 micro edition (J2ME) that aims to serve small devices[19].

J2ME is divided into Configuration, Profile and Optimal Application Programming Interface. Configuration has two categories: Connected Device Configuration (CDC) design for Personal Digital Assistance (PDAs) and limited Connected Device Configuration that is oriented to the mobile device. The profile corresponding to the mobile device in J2ME is called Mobile Information Device Profile(MIDP). The researcher tries to develop an application based on (MIDP) to receive progressive video Stream from HTTP server using the HTTP protocol[1] .

- **HTTP CLIENT:** MIDlet is an application developed according to MIDP in j2me. The MIDlet designed has an important thread called (connection thread). The connection thread is responsible for creating the player and keeping the transmission alive during the playback video stream, while the main program is in charge of preparing the user interfaces and commands. Figure(2) shows the flow chart of the proposed streaming application.

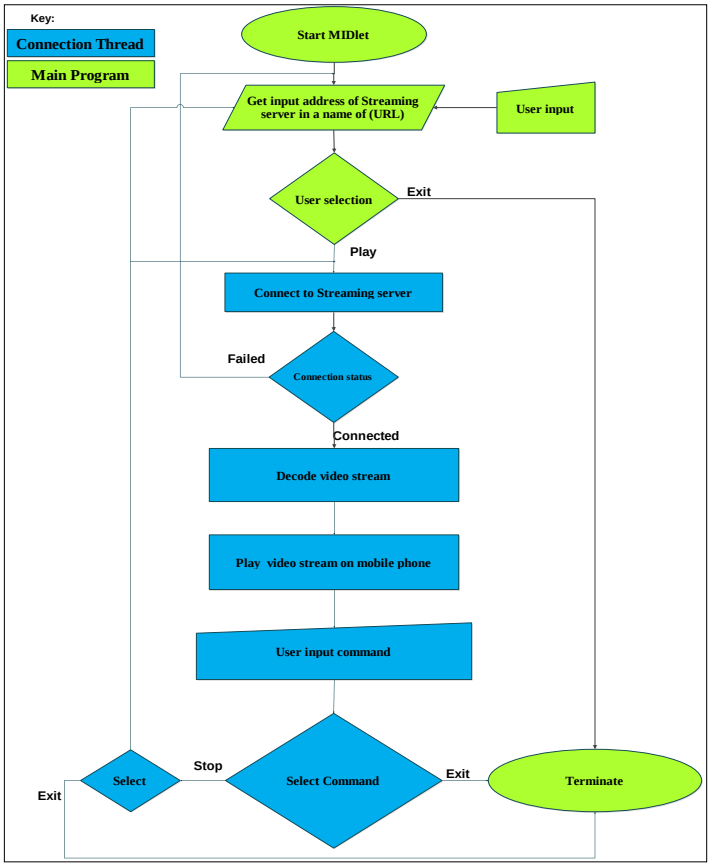


Figure 2: Flow chart of receiving and playback progressive video stream

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Owing to the fact that J2ME does not support the HTTP streaming protocol by itself, it is good idea to utilize the native browser in the mobile phone to use the embedded player that supports the HTTP streaming. After doing this, the mobile phone can be connected to the HTTP server via the HTTP protocol and the player receives and plays the progressive video stream(making buffer to be filled ,then played). Two Special libraries must be included in the j2me produced code to support the HTTP streaming in the mobile phone. These libraries are provided by Nokia corporation to give their products the ability for playing back the progressive video as shown down. Note that Nokia phones have supported progressive playback video since series/40 platform.

package com.nokia.developer.video;
com.nokia.mid.ui

- **MOBILE PHONE OPERATING SYSTEM :**The operating system used in our research is Symbian OS. Most mobile phone manufactures choose Symbian OS for their product since it is designed specifically for the mobile phone. It has a very small memory footprint and a low power consumption. Nowadays, Symbian becomes open OS. Many features make us choose this platform among which are the following: It Supports client -server architecture with a set of API required implementation enabling the third party developers to write and install applications independently from the device manufacturers. Note that the number of devices posed by Nokia in the market in 2007 is (60 million) units [20, 21]. Despite the emergence of many other platforms, this platform has so far been effectively used in the Middle East till this day.

D. *WI-FI CHANNEL*

Recently streaming Audio and video have been popular in wired network, but after the emergence of wireless network, the attention is shifted to delivering video over these networks since they provide flexible connectivity than wired network. In this paper, the transmission of progressive video uses the Wi-Fi channel which can operate at a high bitrate to allow the transmission of high quality video data. WLAN have two major challenges for video streaming: first, changing in the channel quality and second, the high bit error compared with the wired network[22]. Wi-Fi operates in two major modes. The first is infrastructure mode in which the devices are connected via access point, the devices and access point are identified by SSID. The second mode known as the Ad-hoc mode that allows the devices within one another's communication range to Communicate directly without access point[23]. The Ad-hoc mode is used in this paper to make a connection between the mobile device and the HTTP server.

IV. EXPERIMENTS & RESULTS

All experiments have been done on Nokia phones. N86 8MP and C6-01 are the two phones taken under investigation. Each of these phones has its own specifications. Table (1) shows these specifications.

TABLE 1:SMARTPHONES SPECIFICATION USED IN TEST

Smartphone type	CPU speed(MZ)	RAM (MB)	Power Capacity(mAh)
Nokia N86 8MP	484	128	1200
Nokia C6-01	718	256	850

To make the measurements & analysis of progressive video to mobile phone accurate, several points must be taken into account:

- The video file used for the test and measurements is fixed for all experiments represented by (3.96 MB) in size and (1Min) duration before making any CODEC on it just to make sure that all the tests performed on the same video clip have the same properties(frame number , resolution, video size, video duration and video contents). The original video file is downloaded from YouTube under the title "Broadcast Quality Video over Wireless".

- The video file resolution is set to be CIF (320*320) for both MPEG-4 and H.264 CODECs and QCIF(176*144) for H.263 . This disparity in video resolution is because H.263 CODEC supported only QCIF(176*144)[24]. The audio is fixed for all the experiments represented by MP3 with 128 Bit rate and 44100 sample rate 2 channel . Vide Lan Client (VLC) has been used to adjust the video/audio CODECs& the resolution size.

- The signal of the Wi-Fi is assumed to be an excellent signal. This can be proved by making the experiments of the mobile phone while receiving progressive video stream very close to the HTTP server.

- The sound of the test video is disabled (no sound) because the mobile phones have a different sound speaker in terms of volume and power from one to another ,taking into account that sound data will be processed in the mobile phone processing system.

- The brightness of the display screen is a very important element, since it affects the power consumption on the mobile phone during the playback video file. Moreover, the new smartphones nowadays are equipped by their manufactures with the Light-sensitive diode which in turn controls the lighting mobile screen. The measurements adapt a full light on the Light-sensitive diode(Daylight) and adjust the mobile phone display brightness to 75%.

- Smartphones users communicate through 3G networks or any other (CDMA, 4G). This connection also consumes energy from battery. It is not reasonable that the user disconnects his terminal with the 3G network because of his desire to watch a video clip. For this reason the total measurement of power consumption takes into account the 3G network consumption in addition to the playback progressive video file note that the power consumption of the 3G plus OS(standby) in N86 8MP for 1Min is almost 0.15W.and C6-01 is 0.31W.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

In order to complete our analysis of the progressive video streaming on the mobile phone, a special tool is required to measure the CPU and RAM utilization as well as power consumption. The measurements and analysis take place on Nokia devices. The choice of the mentioned commercial devices is made due to several reasons. First, these phones are considered as 3G phones, and secondly, they are able to run an in-built energy profiler developed by Nokia. The Nokia Energy Profiler is an applications for S60 3rd and later additions. It gaves developers information about (power consumption, battery voltage, processor activity, Ram usage and WLAN signal straight, etc.)[25].

A. EXPERIMENTS ON N86 8MP

- **MPEG-4 with N86 8MP**

The CODEC used in N86 8MP is the same as that conduced on C6. Table (2) shows the bitrate with the frame rate conducted on N86 8MP with video CODEC MPEG-4 as well as the CPU & RAM utilization.

TABLE 2 :DIFFERENT SCENARIO OF VIDEO CODEC (MPEG-4) CONDUCTED ON N86 8MP

Video CODEC		Overall utilization		Progressive Video utilization		Overall power Consumption (w	Progressive video consumption(W)	Time can Play progressive due CODEC setting ()
Bitrate kbps	Frame rate	CPU Usage	Memory Usage(MB)	CPU Usage	Memory Usage(MB)			
256	20	49%	65.7	41%	5.1	1.51	1.36	03:13
512	20	50%	66.2	42%	5.6	1.5	1.35	03:14
768	25	55%	64.4	47%	3.8	1.52	1.37	03:10
1024	30	60%	65.2	52%	4.6	1.56	1.41	03:04
1280	30	Mobile phone unable to run both progressive video & Energy profiler						
1536	30							
1792	30							
2048	30							
2304	30							
2560	30							
2816	30							
3072	30							

Figure (3) explains in plot the CPU & RAM activity during the playback video stream with MPEG-4 video CODEC.

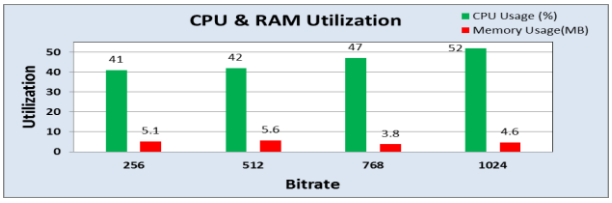


Figure 3: CPU&RAM (N86 8MP, MPEG-4)

Figure (4) shows the power consumption during the play of progressive video stream in N86 using MPEG-4 CODEC, while figure (5) presents in plot the time that the mobile phone can run progressive video encoded in specific configurations above before the mobile battery is exhausted.

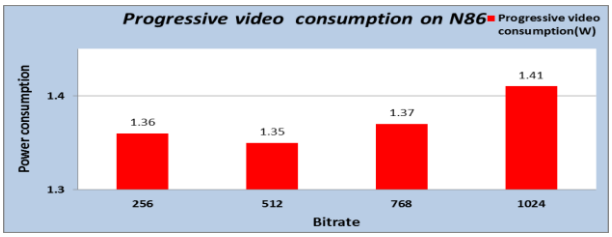


Figure 4: Power consumption (N86 8MP, MPEG-4)

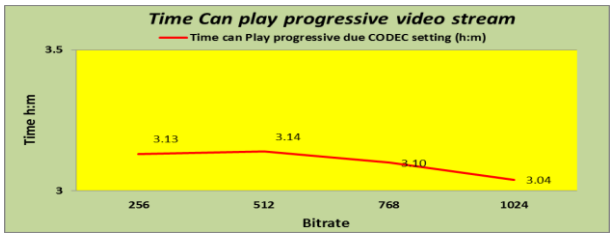


Figure 5: playback progressive video before battery runs out(N86 8MP, MPEG-4)

- **H.263 with N861 8MP**

Table (3) shows the bitrate with the frame rate conducted on N86 8MP with video CODEC H.263 as well as the CPU & RAM utilization.

TABLE 3 :DIFFERENT SCENARIO OF VIDEO CODEC (H.263) CONDUCTED ON N86 8MP

Video CODEC		Overall utilization		Progressive Video utilization		Overall power Consumption (w)	Progressive video consumption(W)	Time can Play progressive due CODEC setting (h:m)
Bitrate kbps	Frame rate	CPU Usage	Memory Usage(MB)	CPU Usage	Memory Usage(MB)			
128	15	43%	68.3	35%	7.7	1.27	1.12	03:48
256	20	46%	68.9	38%	8.6	1.29	1.14	03:44
512	20	46%	69.3	38%	8.7	1.3	1.15	03:42
768	25	48%	69.4	40%	8.8	1.31	1.16	03:40
1024	30	50%	65.2	42%	4.6	1.33	1.18	03:35
1280	30	50%	67.4	42%	6.8	1.36	1.21	03:32
1536	30	50%	67.5	42%	6.9	1.34	1.19	03:33
1792	30	49%	67.4	41%	6.8	1.34	1.19	03:34
2048	30	50%	67.6	42%	7	1.34	1.19	03:34
2304	30	50%	67.7	42%	7.1	1.34	1.19	03:34
2560	30	50%	67.8	42%	7.2	1.36	1.21	03:30
2816	30	51%	67.8	43%	7.2	1.3	1.15	03:39
3072	30	49%	67.7	41%	7.1	1.33	1.18	03:34

Figure (6) explains in plot the CPU & RAM utilization conduced on table (3). The power consumption of N86 during the playback progressive video with H.236 CODEC shown in figure (7).

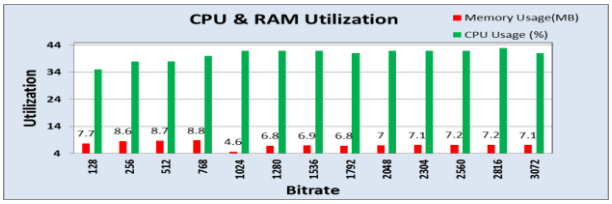


Figure 6: CPU&RAM (N86 8MP, H.263)

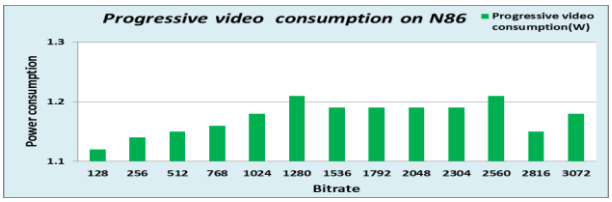


Figure 7: Power consumption (N86 8MP, H.263)

Figure (8) presents in plot the time that mobile phone can run progressive video encoded with H.263 in specific

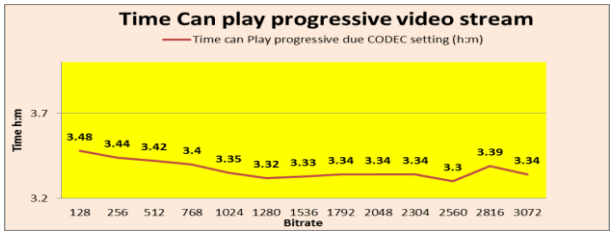


Figure 8: Playback progressive video before battery runs out(N86 8MP, H.263)

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

configurations above(table 3) before mobile battery is exhausted.

- **H.264 with N86 8MP:**

N86 8MP device operates on S60 OS with version (10) does not support this advanced CODEC, but Nokia corporation provides this feature in S60 version (30). Table (4) shows the bitrate with the frame rate conducted on N86 8MP with video CODEC H.264 as well as the CPU & RAM utilization .

TABLE 4 :Different scenario of video CODEC (H.264) conducted on N86 8MP

Video CODEC		Overall utilization		Progressive Video utilization		Overall power	Progressive video	Time can Play progressive
Bitrate kbps	Frame rate	CPU Usage	Memory Usage(MB)	CPU Usage (%)	Memory Usage(MB)	Consumption (w)	consumption(W)	due CODEC setting (h:m)
128	15	45%	67.3	37%	6.7	1.45	1.3	03:19
256	20	48%	67.7	40%	7.1	1.49	1.34	03:13
512	20	51%	67.7	43%	7.1	1.52	1.37	03:09
768	25	54%	68.1	46%	7.5	1.53	1.38	03:06
1024	30	58%	68.3	50%	7.7	1.55	1.4	03:05
1280	30	58%	68.3	50%	7.7	1.57	1.42	03:02
1536	30	Mobile Phone unable to run video with this CODEC configurations						
1792	30							
2048	30							
2304	30							
2560	30							
2816	30							
3072	30							

Figure (9) explains in plot the CPU & RAM utilization conduced on N86 8MP with H.264 video codec.

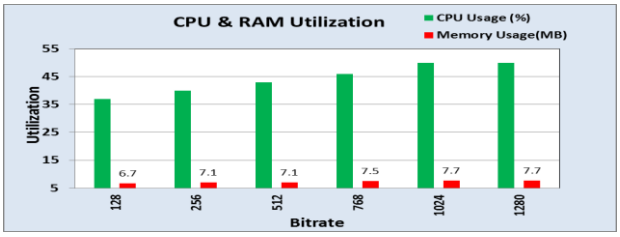


Figure 9: CPU&RAM (N86 8MP, H.264)

Figure (10) shown down presents the time that the mobile phone can run progressive video encoded with H.264 on configurations specified in table (4)above before the mobile battery is exhausted, while figure (11) shown in plot the power consumption in the same setting .

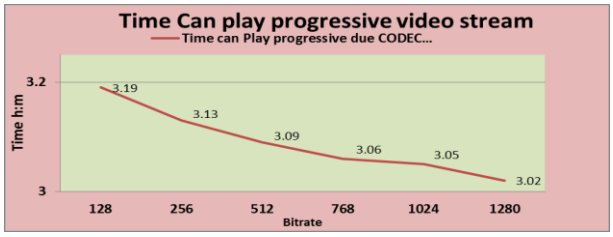


Figure 10: Playback progressive video before battery runs out(N86 8MP, H.264)

B. Experiments on C6_01:

- **MPEG-4 with C6_01:**

Nokia C6-01 has a higher CPU speed than N86 8MP as shown in table (1) above. Table (5) shows the bitrate with the frame rate conducted on C6_01 with video CODEC MPEG-4 as well as the CPU & RAM utilization.

TABLE 5 :Different scenario of video CODEC (MPEG-4) conducted

Video CODEC		Overall utilization		Progressive video utilization		Overall power	Progressive power	Time can Play progressive
Bitrate kbps	Frame rate	CPU usage	Memory usage (MB)	CPU usage	Memory usage (MB)	consumption (W)	consumption (W)	due CODEC setting (h:m)
256	20	45%	125.8	27	13.5	1.17	0.86	02:53
512	20	48%	126.1	30	13.8	1.18	0.87	02:50
768	25	50%	125.9	32	13.6	1.21	0.9	02:45
1024	30	53%	126.1	35	13.8	1.22	0.91	02:43
1280	30	57%	126.7	39	14.4	1.24	0.93	02:40
1536	30	60%	126.8	42	14.5	1.24	0.93	02:39
1792	30	60%	126.5	42	14.2	1.25	0.94	02:38
2048	30	63%	127.3	45	15	1.25	0.94	02:37
2304	30	63%	127.6	45	15.3	1.22	0.91	02:39
2560	30	70%	127	52	14.7	1.29	0.98	02:29
2816	30	70%	127.3	52	15	1.28	0.97	02:30
3072	30	68%	127.5	50	15.2	1.26	0.95	02:31

Figure (12) explains in plot the CPU & RAM utilization conduced on c6-01 with MPEG-4 video codec

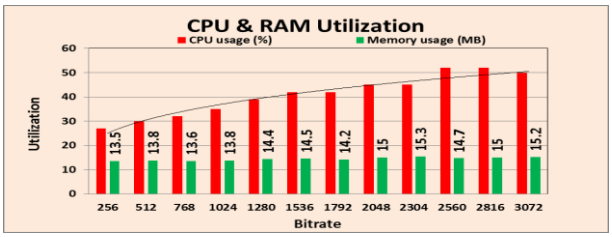


Figure 12: CPU&RAM (C6-01, MPEG-4)

Figure (13) shows the power consumption during the playback progressive video stream on C6-01 with MPEG-4 CODEC.

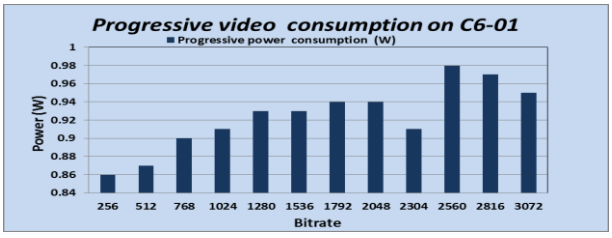


Figure 13: Power consumption utilization (C6-01, MPEG-4)

According to C6-01 battery capacity and power consumption in table (1), the time that the device can play the progressive video befor battery is exhusted shown in figure (14).

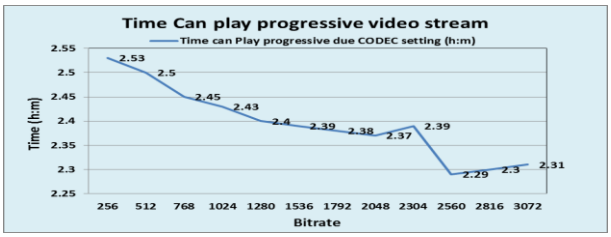


Figure 14: Playback progressive video before battery runs out(C6-01, MPEG-4)

- **H.263 with C6_01:**

Table (6) shows the bit rate with frame rate conducted on C6_01 with video CODEC H.263 as well as the CPU & RAM utilization.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

TABLE 6 :Different scenario of video CODEC (H.263) conducted on C6-01

Video CODEC		Overall utilization		Progressive video utilization		Overall power consumption (W)	Progressive power consumption (W)	Time can Play progressive due CODEC setting (h:m)
Bitrate kbps	Frame rate	CPU usage	Memory usage (MB)	CPU usage	Memory usage (MB)			
128	15	40%	116	22%	3.7	1.16	0.85	02:57
256	20	41%	118.4	23%	6.1	1.18	0.87	02:53
512	20	44%	119.7	26%	7.4	1.18	0.87	02:52
768	25	45%	120	27%	7.7	1.18	0.87	02:52
1024	30	49%	119.9	31%	7.6	1.2	0.89	02:48
1280	30	48%	120.1	30%	7.8	1.21	0.9	02:47
1536	30	48%	120.4	30%	8.1	1.2	0.89	02:48
1792	30	48%	120.2	30%	7.9	1.19	0.88	02:48
2048	30	47%	120.1	29%	7.8	1.2	0.89	02:47
2304	30	48%	120.9	30%	8.6	1.21	0.9	02:45
2560	30	48%	120.9	30%	8.6	1.19	0.88	02:45
2816	30	49%	121.2	31%	8.9	1.2	0.89	02:43
3072	30	49%	121.3	31%	9	1.2	0.89	02:43

The plot shown down in Figure (15) explains Nokia C6-01 with CPU&RAM usage during the playback progressive video stream with a different bitrate using H.263 CODEC.

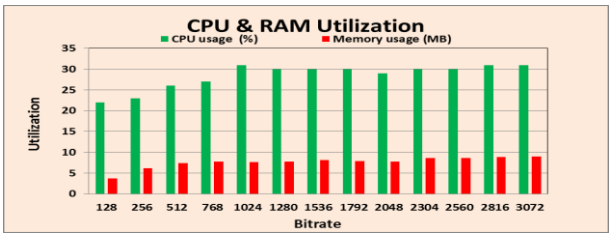


Figure 15: CPU&RAM utilization (C6-01, H.263)

Figure (16) shown down presents the time that mobile phone can run the progressive video encoded with H.263 on configurations specified in table (6) before the mobile battery is exhausted, while figure (17) plots the power consumption in the same setting .

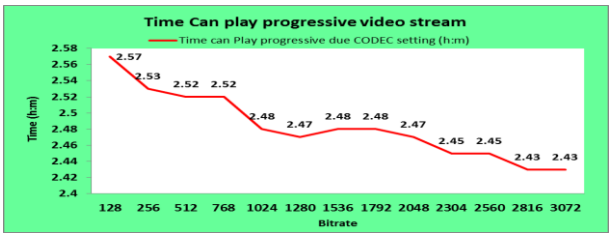


Figure 16: Playback progressive video before battery runs out(N86 8MP, H.263)

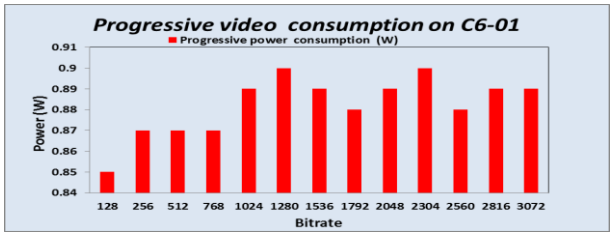


Figure 17: Power consumption utilizations (C6-01, H.263)

- **H.264 vs. C6_01:**

Table (7) shows the bitrate with the frame rate conducted on C6_01 with video CODEC H.264 as well as the CPU & RAM utilization.

TABLE 7 :Different scenario of video CODEC (H.264) conducted on C6-01

Video CODEC		Overall utilization		Progressive video utilization		Overall power consumption (W)	Progressive power consumption (W)	Time can Play progressive due CODEC setting (h:m)
Bitrate kbps	Frame rate	CPU usage	Memory usage (MB)	CPU usage	Memory usage (MB)			
128	15	39%	138.8	21%	26.5	1.11	0.8	02:51
256	20	43%	140	25%	27.7	1.14	0.83	02:46
512	20	45%	140.4	27%	28.1	1.16	0.85	02:43
768	25	49%	138.7	31%	26.4	1.18	0.87	02:40
1024	30	52%	141.2	34%	28.9	1.2	0.89	02:38
1280	30	52%	141.2	34%	28.9	1.19	0.88	02:37
1536	30	51%	141.3	33%	29	1.2	0.89	02:37
1792	30	52%	141.5	34%	29.2	1.2	0.89	02:36
2048	30	54%	141.6	36%	29.3	1.2	0.89	02:35
2304	30	53%	141.7	35%	29.4	1.2	0.89	02:35
2560	30	53%	141.7	35%	29.4	1.2	0.89	02:35
2816	30	53%	141.8	35%	29.5	1.2	0.89	02:34
3072	30	52%	141.8	34%	29.5	1.19	0.88	02:36

Figure (18) explains in plot the CPU & RAM utilization conduced on c6-01 with H.264 video CODEC.

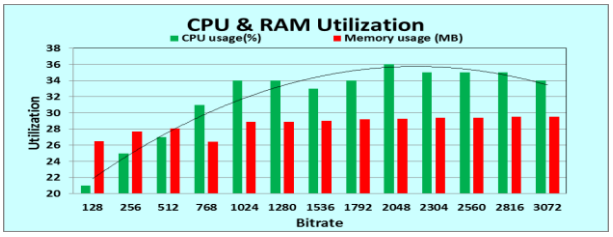


Figure 18: CPU&RAM utilizations (C6-01, H.264)

Figure (19) shown down presents the time that the mobile phone can run progressive video encoded with H.263 on configurations specified in table (7) before the mobile battery is exhausted, while figure (20) plots the power consumption in the same setting.

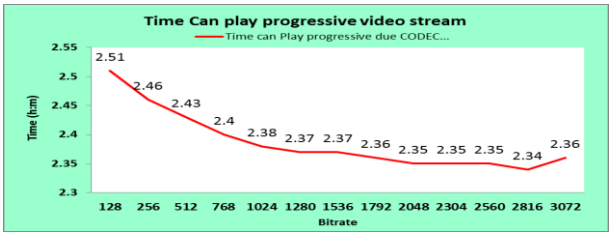


Figure 19: Playback progressive video before battery runs out(C6-01, H.264)

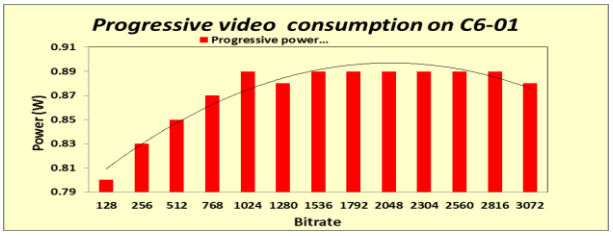


Figure 20: Power consumption utilizations (C6-01, H.264)

V. CONCLUSION

Since mobile phones have different specifications and capabilities. Streaming video to these devices is considered to be a great challenge for the developer. Streaming service require special types of protocols to be handled such as Real

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Time Streaming Protocol (RTSP) and not all mobile phones support this protocol.

This paper presents the implementation of video streaming service called Progressive video stream to those mobile phones that do not support RTSP protocol. This is because the progressive video stream depends on the HTTP protocol which, in turn, considered to be mandatory is included in all mobile phones. J2ME is the programming languages that help to complete this goal through a technology named MMAPI. The channel used is wireless Wi-Fi IEEE802.11. The platform used for work is Symbian. Also the paper presents measurements and analysis of CPU&RAM resources during playback progressive video with different CODEC. A special tool is used in our test named Energy Profiler presented by Nokia Corporation and two mobile phones are used in our test (Nokia N86 8MP and Nokia C6-01) .

From the experiment we conclude that the mobile phone with high processor speed consumes less power than the low processor speed in the same CODEC. This is true because the low CPU mobile phone must increase the operating frequency to meet the performance of the high CPU mobile phone and it is known of the basic semiconductor physics that the increasing operating frequency and voltage can exponentially increase the power consumption of the semiconductor devices[26].

Figure (21) show the comparison in power consumption between N86 and C6-01 in the same CODEC. C6-01 has a processor speed of 718 MHZ while N86 has 484 MHZ. Also the tests show that MPEG-4 CODEC is more CPU usage than other CODEC when playing back the progressive video stream in the same mobile phone. See figure (22).

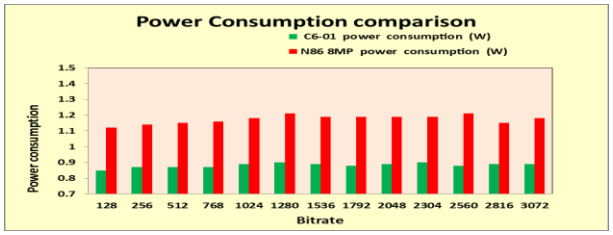


Figure 21:Power consumption Comparison (C60-1 vs. N86 8MP ,H.263)

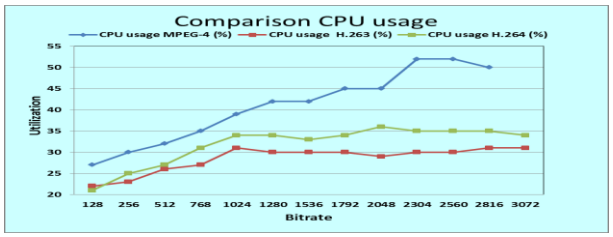


Figure 22:CPU usage Comparison (C60-1, MPEG-4 vs. H.263 vs. H.264)

References

[1] M. Vazquez-Briseno and P. Vincent, "An Adaptable Architecture for Mobile Streaming Applications Summary," *IJCSNS International Journal of Computer Science and Network Security,* vol. VOL.7, pp. 79-84, 2008.

[2] Z.-r. Wang and Z. Liu, "Implementation of Mobile Streaming Media Player Based on BREW *," *Journal of Electronic Science and Technology of China,* vol. Vol.4, pp. 244-248, 2008.

[3] X. Zhang and H. Hassanein, "A survey of peer-to-peer live video streaming schemes - An algorithmic perspective," *Computer Networks,* vol. Vol 56, pp. 3548-3579, 2012.

[4] S. Ary Mazharuddin, P. Henry, and C. Henning Titi, "A Video Streaming Application Using Mobile Media Application Programming Interface," *TELKOMNIKA,* vol. Vol 08, pp. 293-300, 2010.

[5] S. LI and J. KNUDSEN, *Beginning J2ME: From Novice to Professional*, Third Edition ed.: Apress, 2005.

[6] R. Wuling and Y. Dafeng, "Research on encryption technology based on J2ME socket network communication," *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC),* pp. 1969-1973, 2011.

[7] Oracel, "Mobile Media API Version 1.0,Java 2 Platform Micro Edition," june 2002.

[8] A. Fecheyr, "A Review of HTTP Live Streaming," 2010.

[9] Koro, amp, x, A. si, Sze, B. kely, Csa, sza, and A. r, "TrueVod: Streaming or Progressive Downloading?," *IEEE Communications Letters,* vol. Vol 14, pp. 1083-1085, 2010.

[10] W. Dapeng, Y. T. Hou, Z. Wenwu, Z. Ya-Qin, and J. M. Peha, "Streaming video over the Internet: approaches and directions," *Circuits and Systems for Video Technology, IEEE Transactions on,* vol. 11, pp. 282-300, 2001.

[11] W. Simpson, *Video Over IP IPTV, Internet Video, H.264, P2P, Web TV, and Streaming: A Complete Guide to Understanding the Technology*, Second Edition ed. USA: Elsevier, 2008.

[12] W. E. klof, "Adaptive Video Streaming," Master, KTH Information and Communication Technology, 2008.

[13] L. Keller, "Design and Implementation of a Light Mobile Video Streaming Application for Google Android," Department of Informatics, University of Zurich, 2009.

[14] R. Fielding, J. Gettys, J. C. Mogul, H. Frystyk, and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1," 1999.

[15] (2012). *What is the Apache HTTP Server Project*. Available: http://httpd.apache.org

[16] (2012). *NETCRAFT*. Available: http://www.netcraft.com/survey

[17] Nokia, "Video and Streaming in Nokia Phones," Version 1.0; June 16, 2003.

[18] R. Schäfer, T. Wiegand, and H. Schwarz, "The Emerging H.264/AVC Standard," *EBU TECHNICAL REVIEW,* 2003.

[19] J. W. Muchow, *Core J2ME™ Technology & MIDP*: Prentice Hall PTR, 2001.

[20] M. Wei, A. Chandran, H. P. Chang, J. H. Chang, and C. Nichols, *Comprehensive Analysis of SmartPhone OS Capabilities and Performance*, 2009.

[21] O. Oleinicov, M. Hassinen, K. Haataja, and P. Toivanen, "Designing and Implementing a Novel VoIP-Application for Symbian Based Devices," *2009 Fifth International Conference on Wireless and Mobile Communications,* pp. 251-260, 2009.

[22] M. A. Qadeer, R. Ahmad, M. S. Khan, and T. Ahmad, "Real time video streaming over heterogeneous networks," presented at the International Conference on Advanced Communication Technology, 2009.

[23] X. Bo, K. Seada, and N. Venkatasubramanian, "An Experimental Study on Wi-Fi Ad-Hoc Mode for Mobile Device-to-Device Video Delivery," *IEEE INFOCOM Workshops 2009,* pp. 1-6, 2009.

[24] V. Vehkalahti and R. Kantola, "Study of Video Transmission on TETRA Enhanced Data Service Platform ".

[25] B. Wang, J. Kurose, P. Shenoy, and D. Towsley, "A Model for TCP-based Video Streaming."

[26] NIVIDA, "The Benefits of Multiple CPU Cores in Mobile Devices," *Whitepaper,* p. 32, 2010.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

AUTHORS PROFILE

Omar Abdulmunem Ibrahim Al-Dabbagh (PhD) is currently a head of computer and Internet center/ Mosul university and a lecturer at the computer science department, College of Computer Science and Mathematics at Mosul University/ Iraq. He got a Post Doctoral Research Fellow from National Advanced IPv6 Centre of Excellence (NAv6) at Universiti Sains Malaysia (USM)/ Malaysia. Dr. Omar obtained his bachelor, master, and doctorate in computer science from Mosul University in 1998, 2000, and 2006 respectively. His research area include Network protocols, Multimedia Network, Network security and mobile programming.

Ali Hashim Mohammed AL-Shakarchi is currently a master student in computer science at Mosul University. Ali obtained his bachelor in computer science from the Same college in 2003. He Joined at the Ministry of Health / Department of Ninava as a programmer in 2008 . His interested research area include Network protocols, Multimedia communications, Mobile programming, and distributed database.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Knowledge Discovery In Academic Electronic Resources Using Text Mining

Ojo, Adebola K.
Department of Computer Science
University of Ibadan
Ibadan, Nigeria

Adeyemo, Adesesan B.
Department of Computer Science
University of Ibadan
Ibadan, Nigeria

Abstract - Academic resources documents contain important knowledge and research results. They have highly quality information. However, they are lengthy and have much noisy results such that it takes a lot of human efforts to analyse. Text mining could be used to analyse these textual documents and extract useful information from large amount of documents quickly and automatically. In this paper, abstracts of electronic publications from African Journal of Computing and ICTs, an IEEE Nigerian Computer Chapter Publication were analysed using text mining techniques. A text mining model was developed and was used to analyse the abstracts collected. The texts were transformed into structured data in frequency form, cleaned up and the documents split into series of word features (adjectives, verbs, adverbs, nouns) and the necessary words were extracted from the documents. The corpus collected had 1637 words. The word features were then analysed by classifying and clustering them. The text mining model developed is capable of mining texts from academic electronic resources thereby identifying the weak and strong issues in those publications.

Keywords: Text Mining, Academic Journals, Classification, Clustering, Document collection.

1. INTRODUCTION

Text Mining is a process of extracting new, valid, and actionable knowledge dispersed throughout text documents and utilizing this knowledge to better organize information for future reference. Mining implies extracting precious nuggets of ore from otherwise worthless rock [1]. It is the gold hidden in mountains of textual data [2].

Text mining, otherwise known as Text Data Mining (TDM), is the discovery by computer of new, previously unknown information, by automatically extracting information from a usually large amount of different unstructured textual resources. *Previously unknown* implies discovering genuinely new information. *Unstructured* means free naturally occurring texts- as opposed to HyperText Markup Language (HTML), eXtensible Markup Language (XML), and other scripting languages.

Text mining can be described as data mining applied to textual data. Text is "unstructured, amorphous, and difficult to deal with" but also "the most common vehicle for formal exchange of information." [3].

1.1 TDM and Information Retrieval

TDM is a non-traditional information retrieval (IR) whose goal is to reduce the effort required of users to obtain useful information from large computerized text data sources. Traditional IR often simultaneously retrieves both "too little" information and "too much" text [4] [3]. However, in Information Retrieval (Information Access), no genuinely new information is found. The desired information merely coexists with other valid pieces of information.

1.2 TDM, Computational Linguistics and Natural Language Processing (NLP)

If we extrapolate from data mining on numerical data to data mining from text collections, it is discovered that there already exists a field engaged in text data mining: corpus-based computational linguistics! Computational linguistics refers to the long-established interdisciplinary field at the intersection of linguistics, phonetics, computer science, cognitive science, artificial intelligence and formal logic, which again is frequently assisted by statistical techniques [5] [6]. Empirical computational linguistics computes statistics over large text collections in order to discover useful patterns. These patterns are used to inform algorithms for various sub problems within natural language processing, such as part-of-speech tagging and word sense disambiguation [1].

NLP is the branch of linguistics which deals with computational models of language. NLP has several levels of analysis: phonological (speech), morphological (word structure), syntactic (grammar), semantic (meaning of multiword structures, especially sentences), pragmatic (sentence interpretation), discourse (meaning of multi-sentence structures), and world (how general knowledge affects language usage) [7]. When applied to IR, NLP could in principle combine the computational (Boolean, vector space, and probabilistic) models' practicality with the cognitive model's willingness to

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

wrestle with *meaning*. NLP can differentiate *how* words are used such as by sentence parsing and part-of-speech tagging, and thereby might add discriminatory power to statistical text analysis. [3].

1.3 TDM and Data Mining (DM)

In Text Mining, patterns are extracted from natural language text rather than databases. The input is free unstructured text, whilst web sources are structured. Table 2 presents a summarized comparison of Data Mining and Text Data Mining.

Table 2: A Comparison of Data Mining and Text Mining

	DM	TM
Object of Investigation	Numerical and categorical data	Textual Data
Object structure	Structured (Relational database)	Unstructured or Semi-structured (Free form texts)
Goal	Predict outcomes of future situations	Retrieve relevant information, distil the meaning, categorize and target-deliver
Methods	Machine learning: SKAT, DT, NN, GA	Indexing, special neural network processing, linguistics, ontologies
Current market size	100,000 analysts at large and midsize companies	100,000,000 corporate workers and individual users
Maturity	Broad implementation since 1994	Broad implementation starting 2000

The relationship of data mining, information retrieval, statistics, web mining, computational linguistics and natural language processing, to text data mining is shown in Figure 2.

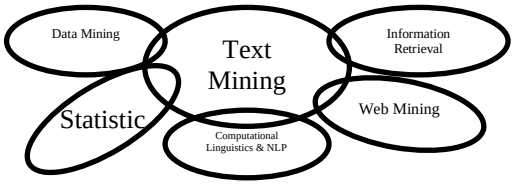


Figure 2: Relationship of Text Mining and Other Applications

2. RELATED WORK

The evolution of internet as a means for sending information led to the growth of on-line knowledge resources and to the diversification of forms and formats used for their storage and transmission: text, data, video and audio. Although hardware restrictions of storage space and data transmission speed is no longer a problem, the text still remains the most efficient form for presenting knowledge over the internet, compared to different audio, video and multimedia formats [8].

With the rapid development of the Internet, the volume of semi-structured and unstructured textual data such as XML documents, e-mail messages, blog posts, academic papers has been under an exponential growth. Discovering useful knowledge from such huge volume of data has become a very challenging problem. Text mining tries to extract knowledge from unstructured data by using techniques from data mining, machine learning, natural language processing, information retrieval, and knowledge management [9]. Text mining is a knowledge-intensive process in which a user interacts with a document collection by a suit of analysis tools, and finally identifies and explores some interesting patterns [9]. Text data mining is a natural extension of data mining [1], and follows steps similar to those in DM. The qualitative difference in text mining, however, is that TDM processes data from natural language text rather than from structured databases of facts [10].

Companies use text mining software to draw out the occurrences and instances of key terms in large blocks of text, such as articles, Web pages, complaint forums, or Internet chat rooms and identify relationships[11]. The software converts the unstructured data formats of articles, complaint forums, or Web pages into topic structures and semantic networks which are important data drilling tools. Often used as a preparatory step for data mining, text mining often translates unstructured text into a useable database-like format suitable for data mining for further and deeper analysis [12]. [13] also described text mining as an emerging technology that can be used to augment existing data in corporate databases by making unstructured text data available for analysis.

[14] classifies text mining techniques into classifier learning, clustering, and topic identification. Classifiers for documents are useful for many applications. Major uses for binary classifiers include spam detection and personalization of streams of news articles. Multiclass classifiers are useful for routing messages to recipients. Most classifiers for documents are designed to categorize according to subject matter. However, it is also possible to learn to categorize according to qualitative criteria such as helpfulness for product reviews submitted by consumers. In many applications of multiclass classification, a single document can belong to more than one category, so it is correct to predict more than one label. This task is specifically called multi-label classification. In standard multiclass classification, the classes are mutually exclusive, that is, a special type of negative correlation is fixed in advance. In multi-label classification, it is important to learn the positive and negative correlations between classes [14]. Another way to view text data mining is as a process of exploratory data analysis that leads to heretofore unknown information, or to answers for questions for which the answer is not currently known. [1]

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Text-mining is ideally suited to extract concepts out of large amounts of text for a meaningful analysis. It has been used in a wide variety of settings, ranging from biomedical applications to marketing and emotional/sentiment research where a lot of data needs to be analyzed in order to extract core concepts. Text-mining achieves this, by applying techniques from information retrieval (such as Google), natural language processing, including speech tagging and grammatical analysis, information extraction, such as term extraction and named-entity recognition and data mining techniques, such as pattern identification [[15] [16].

2.1 Knowledge Management

There is no universally accepted definition of exactly what knowledge is. Some authors define it as the information individuals possess in their minds. This definition is argued by saying that raw data (raw numbers and facts) exist within an organisation. After processing these data they are converted into information and, once it is actively possessed by an individual, this information in turn becomes knowledge. [17] defines knowledge as the justified belief that increases the capacity of an entity to take effective action. Knowledge management is considered as the process of converting the knowledge from the source available to an organisation and then connecting people with that

has been used in a wide variety of settings, ranging from biomedical applications to marketing and emotional/sentiment research where a lot of data needs to be analyzed in order to extract core concepts. Text-mining achieves this, by applying techniques from information retrieval (such as

Applications of text mining methods are diverse and include Bioinformatics [27], Customer profile analysis, Trend analysis, Anti-Spam Filtering of Emails, Event tracks, Text Classification for News Agencies, Web Search and Patent Analysis [27].

Applications of text mining can also extend to any sector where text documents exist. For instance, history and sociology researchers can benefit from the discovery of repeated patterns and links between events, *crime detection* can profit by the identification of similarities between one crime and source of a *news article* [32] and monitoring inconsistencies between *databases and literature.* [33]. [34] presents the framework of the proposed work.

knowledge. The aim of knowledge management is the *creation*, *access* and *reuse* of knowledge [17].
Traditionally, textual elements are extracted and applied in the data mining phase aiming to reveal useful patterns [18]. [19] concentrated on the extraction of textual elements (that is, entities and concepts). Thus the extraction and correlation of textual elements are the basis for the data mining and information retrieval phases aiming to promote support to knowledge management applications.
Knowledge management is seen as systematic and disciplined actions in which organisation can take advantage to get some return [20]. According to [21], knowledge management is an important tool for the documents may be used in order to populate and update scientific database [29]. Other areas include updating automatically a *calendar* by extracting data from *e-mails* [30], [31], identifying the original enhancement of the organisational knowledge infrastructure. The information technology has an important role in the process of transformation of the knowledge, from ***tacit*** to ***explicit*** [22]. Thus we state making explicit entities and their relationships through information extraction and retrieval, and text mining techniques is an important step towards knowledge management applications, such as, communities of practice [23], [24], expertise location [22] and competency management [25], [26].
Text-mining is ideally suited to extract concepts out of large amounts of text for a meaningful analysis. It Google), natural language processing, including speech tagging and grammatical analysis, information extraction, such as term extraction and named-entity recognition and data mining techniques, such as pattern identification [15] [16].

3. METHODOLOGY

The overall process of conducting text-mining-based analysis goes through several steps. This is depicted in Figure 3 below. First of all, text collection and text pre-processing are the preliminary steps.

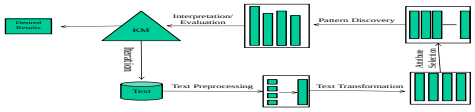


Figure 3: Text Mining Process

Second, raw journal article documents are transformed into structured data. In relation to this analysis, text mining is used as a data processing and information-extracting tool. For mining document

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

collections the text documents are pre-processed and the information stored in a data structure. A text document can be represented by a set of words, that is, a text document is described based on the set of words contained in it (bag-of-words representation). However, in order to be able to define at least the importance of a word within a given document, usually a vector representation is based, where for each word a numerical "importance" value is stored.

Text Pre-processing

In order to obtain all words that are used in a given text, a *tokenization* process is required, that is, a text document is split into a stream of words by removing all punctuation marks and by replacing tabs and other non-text characters by single white spaces. This tokenized representation is then used for further processing. The set of different words obtained by merging all text documents of a collection is called the *dictionary* of a document collection.

In order to allow a more formal description of the algorithms, we define some terms and variables that will be frequently used in the following: Let D be the set of documents and $T= \{t_1, \ldots, t_m\}$ be the dictionary, that is, the set of all different terms occurring in D, then the absolute frequency of term $t \in T$ in document $d \in D$ is given by tf(d,t). We denote the term vectors $\vec{t} = \left(tf(d,t_1), \ldots, tf(d,t_1)\right)$. We also need the notion of the centroid of a set X of term vectors. It is defined as the mean value $\vec{t}X := \frac{1}{|X|}\sum_{\vec{t_d} \in X} \vec{t}d$ of its term vectors. In the sequel, we will apply *tf* also on subsets of terms: For $T' \subseteq$ T, we let ft(d, T'):= $\sum_{t \in T'} tf(d,t)$

Text Transformation and feature selection

In order to reduce the size of the dictionary and thus the dimensionality of the description of documents within the collection, the set of words describing the documents can be reduced by filtering and lemmatization or stemming methods.

Filtering, Lemmatization and Stemming

Filtering methods remove words from the dictionary and thus from the documents. A standard filtering method is stop word filtering. The idea of stop word filtering is to remove words that bear little or no content information, like articles, conjunctions, prepositions. Furthermore, words that occur very seldom are likely to be of no particular statistical relevance and can be removed from the dictionary [27]. In order to further reduce the number of words in the dictionary, also (index) term selection methods can be used.

Lemmatization methods try to map verb forms to the infinite tense and nouns to the singular form. However, in order to achieve this, the word from has to be known, that is, the part of speech of every word in the text document has to be assigned. Since this tagging process is usually quite time consuming and still error-prone, in practice frequently stemming methods are applied.

Stemming methods try to build the basic forms of words, that is, strip the plural 's' from nouns, them 'ing' from verbs, or other affixes. A stem is a natural group of words with equal (or very similar) meaning. After the stemming process, every word is represented by its stem. A well-known rule based stemming algorithm has been originally proposed by Porter (1980). He defined a set of production rules to iteratively transform (English) words into their stems.

Index Term Selection

To further decrease the number of words that should be used also indexing or keyword selection algorithms can be used. In this case, only the selected keywords are used to describe the documents. A simple method for keyword selection is to extract keywords based on their entropy. For each word t in the vocabulary the entropy can be can be computed as

$$W(t) = 1 + \frac{1}{log_2|D|}\sum_{d \in D} P(d,t) log_2 P(d,t) \text{ with}$$

$$P(d,t) = \frac{tf(d,t)}{\sum_{l=1}^{n} tf(d_l,t)} \quad (2)$$

Here the entropy gives a measure how well a word is suited to separated documents by keyword search. Words that occur in many documents will have low entropy. The entropy can be used as a measure of the importance of a word in the given domain context. As index words a number of words that have a high entropy relative to their overall frequency can be chosen, that is, of words occurring equally often those with the higher entropy can be preferred.

In order to obtain a fixed number of index terms that appropriately cover the documents, a simple greedy strategy is applied: From the first document in the collection we select the term with the highest relative entropy as an index term. Then we mark this document and all other documents containing this term. From the first of the remaining unmarked documents we select again the term with the highest relative entropy as an index term. We then mark again this document and all other documents containing this term. We repeat this process until all documents are marked, and then we unmark them all and start again. The process can be terminated when the desired number of index terms has been selected.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The Vector Space Model

Despite of its simple data structure without using any explicit semantic information, the vector space model enables very efficient analysis of huge document collections.

The vector space model represents documents as vectors in *m*-dimensional space, that is, each document *d* is described by a numerical feature vector $w(d) = (x(d,t_1), \ldots, (x(d,t_m))$. Thus, documents can be compared by use of simple vector operations and even queries can be performed by encoding the query terms similar to the documents in a query vector. The query vector can then be compared to each document and a result list can be obtained by ordering the documents according to the computed similarity [27]. The main task of the vector space representation of documents is to find an appropriate encoding of the feature vector.

Each element of the vector usually represents a word (or a group of words) of the document collection, that is, the size of the vector is defined by the number of words (or groups of words) of the complete document collection. The simplest way of document chances of being retrieved independent of their lengths:

$$w(d,t) = \frac{tf(d,t)\log\left(\frac{N}{n_t}\right)}{\sqrt{\sum_{j=1}^{m} tf(d,t_j)^2 \left(\log\left(\frac{N}{n_{t_j}}\right)\right)^2}}, \qquad (3)$$

Where N is the size of the document collection D and n_t is the number of documents in D that contain term *t*.

Based on a weighting scheme a document *d* is defined by a vector of term weights $w(d)=(w(d,t1), \ldots,$ A frequently used distance measure is the Euclidian distance. We calculate the distance between two text documents $d_1, d_2 \in D$ as follows:

$$dist(d_1,d_2) = \sqrt[2]{\sum_{k=1}^{m} |w(d_1,t_k) - w(d_2,t_k)|^2} \qquad (5)$$

However, the Euclidean distance should only be used for normalized vectors, since otherwise the different lengths of documents can result in a smaller distance between documents that share less words than between documents that have more words in common and should be considered therefore as more similar.

For normalized vectors the scalar product is not much different in behaviour from the Euclidean distance, since for two vectors $\vec{x}$ and $\vec{y}$ it is

$$cos\varphi = \frac{\vec{x}\vec{y}}{|\vec{x}||\vec{y}|} = 1 - \frac{1}{2} d^2\left(\frac{\vec{x}}{|\vec{x}|}, \frac{\vec{y}}{|\vec{y}|}\right). \qquad (6)$$

encoding is to use binary term vectors, that is, a vector element is set to one of the corresponding word is used in the document and to zero if the word is not. This encoding will result in a simple Boolean comparison or search if a query is encoded in a vector. Using Boolean encoding the importance of all terms for a specific query or comparison is considered as similar. To improve the performance usually term weighting schemes are used, where the weights reflect the importance of a word in a specific document of the considered collection. Large weights are assigned to terms that are used frequently in relevant documents but rarely in the whole document collection (Hotho, et al 2005). Thus a weight *w(d,t)* for a term *t* in document *d* is computed by term frequency *tf(d,t)* times inverse document frequency *idf(t)*, which describes the term specificity within the document collection. In Salton, et al (1994) a weighting scheme was proposed that has meanwhile proven its usability in practice. Besides term frequency and inverse document frequency – defined as $idf(t) := log\frac{N}{n_t}$ -, a length normalization factor is used to ensure that all documents have equal $(w(d,t_m))$ and the similarity *S* of two documents d_1 and d_2 (or the similarity of a document and a query vector) can be computed based on the inner product of the vectors (by which – if we assume normalized vectors – the cosine between the two document vectors is computed), that is,

$$S(d_1,d_2) = \sum_{k=1}^{m} w(d_1,t_k).w(d_2,t_k). \qquad (4)$$

Part-of-speech tagging (POS) determines the part of speech tag, for example, noun, verb and adjective for each term.

Text chunking aims at grouping adjacent words in a sentence. An example of a chunk is the noun phrase "the current account deficit".

Word Sense Disambiguation (WSD) tries to resolve the ambiguity in the meaning of single words or phrases. An example is 'bank' which have - among others – the senses 'financial institution' or the 'border of a river or lake'. Thus, instead of terms the specific meanings could be stored in the vector space representation. This leads to a bigger dictionary but considers the semantic of a term in the representation.

Parsing: This produces a full parse tree of a sentence. From the parse, we find the relation of each word in the sentence to all the others, and typically also its function in the sentence (for example, subject, object).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The algorithm [35] for text extraction is given as:

{
1 Convert the text into a LIST of words
2 Set threshold to a certain value such as 1 or 2, put a separator to the end of LIST and Set an array LIST[N], an array FinaList[N]=0,
3 Do
{
3.1 Set the frequency of the separator (separator=0)
3.2 Set MergerList[N]=0,
3.3 For i from 1 to NumOf(LIST) – 1 step 1
{
3.4 If LIST[i] is the separator, then Go to Label 3.3.
3.5If Freq(LIST[i]>threshold and
Freq(LIST[i+1] > threshold, then
Merge LIST[i] and LIST[i+1] into MergeList
Else
If Freq(LIST[i])> threshold LIST[i] did not merge with LIST[i-1], then
Save LIST[i] into FinaList.
If the last element of MergeList is not the separator, then
Put the separator to the end of MergeList.
}
4 Set MergeList to LIST
}while NumOf(LIST) <2
5 Filter terms in FinaList
}

4. Results and Discussion

Document Collection

This involves the gathering of academic journal articles using academic electronic resources from African Journal of Computer and ICT, IEEE Nigerian Section.

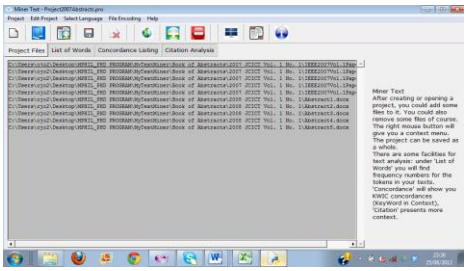


Figure 4: Document Collection

Text Extraction: This involves the identification and extraction of texts from those scientific publications. These raw article documents are then transformed into structured data as shown in Figure 5 below:

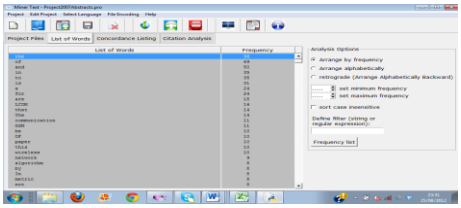


Figure 5: Text Extraction

THE CLUSTERING RESULTS: Overview of the Data

Keywords:

- Data Communication (D):
Broadcast, Radio, acoustic, transmitters, receivers (5)
- **Technology/ICT (T):**
Hardware, Software, Storage device, Coding, Computers, Electronics (7)
- **Location (L):**
world, country, Nigeria (3)
- **Field/Discipline (F):**
Science, Education, Engineering, Medical (4)
- **Product/Market (P):**
result, expansion, advertiser, advancement, economy, present, exploration, finances (8)
- **Organisation (O):**
Government, professionals, subscribers, entrepreneurship (4)
- **Papers/Journal (J):**
published, research, scholars, review (4)
- **Unit (U):**
Age, number, year (3)
- **Facility (Y):**
BCOS, NTA, AIT, Channel, television (5)
- **Method (M):**
Approaches, Measures, techniques, factors (4)
- **Person (N):**
Noble, group, we, I (4)
- **Miscellaneous (S):** other words which did not fall into any of the categories above.

(The numbers in the parenthesis indicate the total number of keywords used during text search.)

4.1 Text Pre-Processing, Transformation and Feature Selection

These involve Text Clean up and tokenization. The document is split into a series of words (features). Stop Words were removed, and words stemmed down to their roots.

4.2 Attribute Generation

Attributes generated are merely labels of the classes automatically produced by a classifier on the features that passed the feature selection process. After this, the database is populated as a result of the process above.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table 3: Attribute Generation

ABSTRACT	DATA COMM	TECHN	LOCATIO	FIELD/DISC	PRODUCT	ORGAN	PAPERS	UNIT	FACILITY	METHOD	PERSON	MISCEL	STOP WO	TOTAL
1	0	20	19	4	27	1	21	9	0	0	0	53	142	296
2	3	8	8	7	14	3	0	3	0	0	1	27	45	119
3	36	18	1	0	26	0	1	4	0	1	5	18	55	165
4	25	4	8	0	29	0	0	5	17	0	1	15	75	179
5	9	16	0	2	18	1	1	12	0	19	0	13	59	150
6	28	25	0	0	2	0	1	11	0	15	4	27	68	181
7	0	3	1	0	9	0	0	7	0	6	1	12	56	95
8	34	2	4	0	9	0	1	2	1	18	0	21	59	151
9	0	2	1	0	7	0	2	5	0	39	1	22	45	124
10	0	2	0	0	13	0	0	3	0	51	24	11	73	177
TOTAL	135	100	42	13	154	5	27	61	18	149	37	219	677	1637

From Table 3, the corpus consists of abstracts taken from the journal articles (as a sample), having a total number of 1637 words including keywords, title words, and the clue words. The rest are stop words. The keywords, title words and the clue words are all categorised as Data Communications (e.g. transmitters, receivers, bandwidth, broadcast, radio link), Technology/ICT(e.g. software, hardware, devices, computers), Location (e.g., world, Nigeria, Africa, country), Field/Discipline (e.g. Science, Education, Engineering), Product/Market (result, economy, expansion), Organisation (Government, entrepreneurship, professionals), Papers/Journals (research, review, published), Unit, Facility (age, number, year), Methods (approaches, techniques, algorithms, measures), Person (person, noble, group), and Miscellaneous (e.g. used, suggests, offers). Stop words include words such as 'the', 'is', 'of', and 'to'.

Table 4: Attribute Selection

ABSTRACT	DATA COMM	TECHN	LOCATIO	FIELD/DISC	PRODUCT	ORGAN	PAPERS	UNIT	FACILITY	METHOD	PERSON	MISCEL	STOP WO
1	0	4	4	1	6	1	5	2	0	0	0	3	8
2	1	2	2	2	3	1	0	1	0	0	1	2	3
3	8	4	1	0	6	0	1	1	0	1	1	1	3
4	5	1	2	0	6	0	0	1	4	0	1	1	4
5	2	4	0	1	4	1	1	3	0	4	0	1	3
6	6	5	0	0	1	0	1	3	0	3	1	2	4
7	0	1	1	0	2	0	0	2	0	2	1	1	3
8	7	1	1	0	2	0	1	1	1	4	0	2	3
9	0	1	1	0	2	0	1	1	0	8	1	2	3
10	0	1	0	0	3	0	0	1	0	11	5	1	4

Table 4 was generated from Table 3 using the following class intervals: 1 (1-5), 2 (6-10), 3(11-15), 4(16-20), 5(21-25), 6(26-30), 7(31-35), 8(36-40); and for miscellaneous data and stop words, the following class intervals: 1(1-20), 2(21-40), 3(41-60), 4(61-80), 5(81-100); 6(101-120), 7(121-140), 8(141-160), and 9(161-180). This is to reduce the population of data.

By taking each attribute as an effect, Probability Models were generated from Table 4, by taking Probability $Pr = \frac{n(e)}{N}$. The resulting output was given in Table 5.

Table 5: Probability of Occurrence Of Each Attribute

EVENT	DATA COMM.	TECH/ICT	LOCATIO	FIELD/DIS	PRO/MKT	ORG	PAP/JO	UNIT	FACILITY	METHOD	PERSON	MISCE	STOP WOR
RAW DATA (N=10)	0	1	1	1	1	1	1	0	0	0	0	1	1
	1	1	1	1	1	1	0	1	0	0	1	1	1
	1	1	1	0	1	0	1	1	0	1	1	1	1
	1	1	1	0	1	0	0	1	1	0	1	1	1
	1	1	0	1	1	1	1	1	0	1	0	1	1
	1	1	0	0	1	0	1	1	0	1	1	1	1
	0	1	1	0	1	0	0	1	0	1	1	1	1
	1	1	1	0	1	0	1	1	1	1	0	1	1
	0	1	1	0	1	0	1	1	0	1	1	1	1
	0	1	0	0	1	0	0	1	0	1	1	1	1
TOTAL	6	10	7	3	10	3	6	10	2	7	7	10	10
PROB. OF OCCURRENCE	0.6	1	0.7	0.3	1	0.3	0.6	1	0.2	0.7	0.7	1	1

In Table 5, each attribute is taken as an event. When an event occurs, the attribute is assigned 1; otherwise, it is assigned zero (0). It is observed from the above that probabilities of data in Groups Technology/ICT and Product/Market are one (1). This means that most of these journals concentrated on the category Technology/ICT, which involves the use of hardware, software, devices, computers and electronics.

Furthermore, it was discovered that stop words had the highest frequency in the whole corpus. After filtering, there was more concentration on Products/Market, and Methods used. This is further represented graphically in Figures 6, 7 and 8:

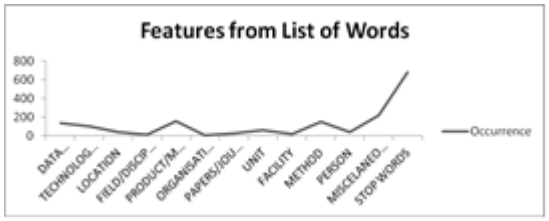


Figure 6: All Attributes Considered

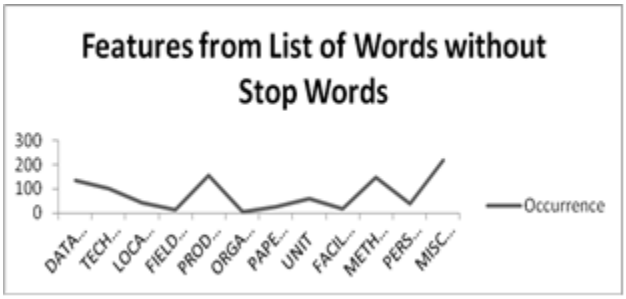


Figure 7: All Attributes Without Stop Words

Figure 8: All Attributes Without Stop Words and Miscellaneous

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Table 6: Correlations among the Attributes

	Data Communication	ICT	Location	Field & Discipline	Product &Market	Organisation
Data Communication	1	.069	-.768	-1.000**	.032	-1.000**
I CT	.069	1	.546	-.737	.161	-.945
Location	-.768	.546	1	-1.000**	.551	-1.000**
Field and Discipline	-1.000**	-.737	-1.000**	1	-.408	.918
Product and Market	.032	.161	.551	-.408	1	-.737
Organisation	-1.000**	-.945	-1.000**	.918	-.737	1

Table 6 shows the correlations (relationships) among all the attributes. It was discovered that there were correlations between some attributes: between attributes ICT and Location (0.546) where ICT was the dependent variable while Location was independent; Product and Location (0.551) where Product was a dependent variable while Location was independent; Paper/Journal and Methods Used (0.847) where former was a dependent variable while the latter was the independent one.

5. Conclusion

Academic resources documents contain important knowledge and research results. They have highly quality information. However, they are lengthy and have much noisy results such that it takes a lot of human efforts for analysis. Text mining could be used to analyse these textual documents and extract useful information from large amount documents quickly and automatically.

This study provides a method for analysing unstructured text. The software captures some selected abstracts of academic publications from the universities electronic resources websites. The processed data was then 'mined' to identify patterns and extract valuable information and new knowledge. The study revealed some strong areas of focus by the authors of these articles in this journal while less concentration was on other areas. This will enable us to have a greater understanding of the patterns and trends of data in these journal articles in future. It will be useful to shape the debate about future research and publications, and hopefully engage current authors of these articles to go beyond the most published (Data Communications) and into other areas of applications.

This research work was based on the academic resources in a particular journal which was specifically based on Data Communications. It can thus be extended to cater for all the journal articles, which cut across other disciplines and fields in Computer Science, as well as all other areas and disciplines in the academic world. This will enable us to know the trends of those publications when taken periodically. Furthermore, it can also be extended to texts being generated by business, academic and social activities – in for example competitor reports, research publications, or customer opinions on social networking sites to capture knowledge and trends.

RERERENCES

[1] M. Hearst, (1999) "Untangling Text Data Mining," in the *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*.

[2] Dorre, J., Gersl, P., & Seiffert, R. 1999. Text mining: Finding nuggets in mountains of textual data. (KDD-99, Association of Computing Machinery, 8, 223-239.

[3] Sharp, M. 2001. Text Mining. Term Paper in Information Studies. Rutgers University, School of Communication, Information and Library Studies. 11 December 2001

[4] Humphreys, K., Demetriou, G., & Gaizauskas, R. 2000. Bioinformatics applications of information extraction for scientific journal articles. Journal of Information Science, 26, 75-85.

[5] Clegg, A. B. 2008. Computational-Linguistic Approaches to Biological Text Mining. A PhD Thesis submitted to the

School of Crystallography, Birkbeck, University of London, Malet Street, UK.

[6] Jurafsky, D. and Martin, J. H. 2000. Speech and Language Processing. Prentice Hall, New Jersey.

[7] Bird, S., Klein, E., and Loper, E. 2007. Natural Language Processing in Python. Draft Copy. University of Pennsylvania. October 12, 2007

[8] Vespan, Dragos M., (2009). PhD Thesis Review: Knowledge Acquisition through Text Mining. *Informatica Economică*, vol. 13, no. 2/2009

[9] L. Jing and R. Y. K. Lau (2009). Granular Computing for Text Mining: New Research Challenges and Opportunities. SpringerLink. Abstract.

[10] Kuan C. Chen, (2009). Text Mining e-Complaints Data From e-Auction Store With Implications For Internet Marketing Research Purdue University Calumet, USA. Journal of Business & Economics Research – May, 2009 Volume 7, Number 5

[11] Robb, Drew. Taming Text. (2005),

[12] Cerrito, Patricia. Inside Text Mining. March 24, 2005

[13] Louise Francis and Matt Flynn. (2010) Text Mining Handbook. Casualty Actuarial Society *E-Forum*

[14] Elkan, C. 2011. Text Mining and Topic Models. elkan@cssd.edu

[15] JISC, 2008. Text Mining Briefing Paper, Joint Information Systems Committee, accessed from

[16] Dahl, Stephan (2010) 'Current Themes in Social Marketing Research: Text-Mining the Past Five Years', Social Marketing Quarterly, 16: 2, 128 — 136

[17] Nonaka, I., von Krogh, G. 2009. "Tacit Knowledge and Knowledge Conversion: Controversy and Advancement in Organizational Knowledge Creation Theory". *Organization Science* **20** (3): 635–652. doi:10.1287/orsc.1080.0412.

18] Mooney, R. J. and Nahm, Un Y. (2005) "Text Mining with Information Extraction". In: Proceedings of the 4th International MIDP colloquium, September 2003, Bloemfontein, South Africa, Daelemans, W., du Plessis, T., Snyman, C. and Teck, L. (Eds.), Van Schaik Pub., South Africa, p. 141-160, 2005.

[19] Goncalves, A. L., Beppler, F. Bovo, A., Kern, V. and Pacheco, R. 2006. A Text Mining Approach Towards Knowledge Management Applications.

[20] Davenport, T. H. and Prusak, L. (1997). "Information ecology: Mastering the information and knowledge environment", Oxford University Press.

[21] Schreiber, G., Akkermans, H., Anjewierden, A., Hoog, R. de, Shadbolt, N., Velde, W. V. de and Wielinga, B. (2002), Knowledge engineering and management: The CommomKADS Metlhodology, The MIT Press, 3rd edition.

[22] Marwick, A.D. (2001) "Knowledge management technology". IBM Systems Journal, v. 40, n. 4, p. 814-830.

[23] Lesser, E. L. and Storck, J. (2001) "Communities of practice and organizational performance", IBM Systems Journal, v. 40, n. 4, p. 831-841.

[24] Wenger E. (1998), Communities of practice, learning meaning and identity, Cambridge University Press, Cambridge, MA.

[25] Dawson, K. (1991) "Core competency management in R&D organizations", In *Technology Management: The New International Language*, Dundar Kocaoglu and Kiyoshi Niwa (eds.), New York, Institute of Electrical and Electronics Engineers, p. 145-148.

[26] Hafeez, K., Zhang, Y. and Malak, N. (2002) "Identifying core competence", IEEE Potencials, v.
49, n. 1, p. 2-8.

[27] Hotho, A., Nurnberger, A. and Paaß, G. 2005. A Brief Survey of Text Mining. Retrieved on April 4, 2011.

[28] Fan, W., Wallace, L., Rich, S. and Zhang, Z. 2006. Tapping the power of text mining. In Communications of the ACM 49(9), pp. 76-82.

[29] Swanson, D. R. and Smalheiser, N. R. 1997. An interactive system for finding complementary literatures: a stimulus to scientific discovery. Artificial Intelligence 91, pp. 183 – 203.

[30] K. Nigam, A. McCallum, S. Thrun, and T. Mitchell, (2000) "Text Classification from Labeled and Unlabeled Documents using EM," in *Machine Learning*, 2000.

[31] Stavrianou, A., Andritsos, P., and Nicoloyannis, N.2007. Overview and Semantic Issues of Text Mining. SIGMOD Record, September 2007 (Vol. 36, No. 3).

[32] Metzler, D., Bernstein Y., Croft, W. B., Moffat, A. and Zobel, J. 2005. Similarity

measures for tracking information flow. In Proc. Of CIKM, Bremen, Germany, pp. 517-524.

[33] Nenadic, G. and Ananiadou, S. 2006. Mining semantically related terms from biomedical literature. In ACM TALIP Special Issue on text Mining and Management in Biomedicine, 5(1), pp 22-43.

[34] Ojo, A. K. and Adeyemo, B. A. (2012). A Framework for Knowledge Discovery from Journal Articles Using Text Mining Techniques. IEEE Journal… Ojo, A. K. & Adeyemo, A. B. (2012): "A Framework for Knowledge Discovery from Journal Articles Using Text Mining Techniques". African Journal of Computing & ICTs (An IEEE Nigeria Computer Chapter Publication) Vol. 5, No. 2, March, 2012 33-42 www.ajocict.net

[35] Liang Yanhong, Tan Runhua. A Text-Mining-Based Patient Analysis in Product Innovative Process. Hebei University of Technology

AUTHORS PROFILE

Dr. Adesesan Barnabas ADEYEMO is a Senior Lecturer at the Computer Science Department of the University of Ibadan. He obtained his PhD, M. Tech., and PGD Computer Science degrees at the Federal University of Technology, Akure. His research interests are in Data Mining, Data Warehousing & Computer Networking. He is a member of the Nigerian Computer Society and the Computer Professionals Registration Council of Nigeria. Dr Adeyemo is a Computer Systems and Network Administration Specialist with expertise in Data Analysis and Data Management.

Adebola K. OJO is a lecturer in the Department of Computer Science, University of Ibadan, Nigeria. She is a registered member of the Computer Professional of Nigeria (CPN). She had her Masters of Science Degree in Computer Science from University of Ibadan, Nigeria. Her research interests are in Digital Computer Networks, Data Mining, Text Mining and Computer Simulation. She is also into data warehouse architecture, design and data quality via data mining approach.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A Comparative Evaluation of Security Aspects of VoIP Technology

[1]Mohd Rahul [2]Mohd Asadullah[3]Md Shabbir Hassan [4]Mohd Muntjir [5]Ahmad Tasnim Siddiqui
College of Computers and Information Technology, Taif University
Saudi Arabia

***Abstract*— Voice over IP (VoIP) technology is swiftly accepted by consumers, militaries, enterprises and governments. This technology recommend higher flexibility and more features than traditional telephony (PSTN) infrastructures, over and above the potential for lower cost through equipment consolidation, new business models for the consumer market. Voice over IP (VoIP) communications is becoming essential to the corporate world. Possibly, Voice over IP should be viewed as a chance to develop new, more effective security policies, infrastructure and processes. These all new policies and practices can have a positive impact on the security of the entire network not only voice communications. This paper provide starting point for understanding the security facets of VoIP in a rapidly evolving set of technologies that are seeing growing deployment and use. The main goal is to provide a better understanding of the security background with respect to VoIP security facet toward directing future research and in other similar up-and-coming technologies.**

Keywords*— *VoIP, ITU-T H.323, Session Initiation Protocol, Media Gateway Control Protocol, Security attacks.

I. INTRODUCTION

In VoIP technology, VoIP is a technology for producing telephone services on IP-based networks. Usually, public switched telephone network (PSTN/ISDN) provides these telephone services, which has been managed and completely controlled by singles, national telephone operators in each country. The voice signal is first divided into frames, then stored in data packets, and lastly transported over internet protocol network using voice communication protocols. Presently, most VoIP systems use either one of two standards; H.3231 or the SIP (Session Initiation Protocol) [1].

VoIP produced a lot of excitement towards the end of the 90s, with the guarantee of providing a possible technology for the journey from the monolithic public switched telephone network (PSTN/ISDN) to next generation networks for which telephone services are produced on an IP-based network. At the turn of the millennium, it was announced that the IETF's Session Initiation Protocol (SIP) standard would be selected as the basis for the 3GPP IP multimedia subsystem (IMS). SIP at this point, was still in an early phase of development. Problems with poor voice quality for the early Internet-based offerings, along with the added barrier of cumbersome technology, e.g., having to phone from the PC made it difficult for consumers to embrace the new technology, and result to slow adoption rate.

The immaturity of the up-and-coming SIP standard contributed mostly to the slowdown of the roll out of VoIP services along with insecurity in the economic and market related factors, and the lack of a solid business model. Today, VoIP is being used all over the place with different levels of success. Home users may use an Analogue Terminal Adapter (ATA) to use their legacy POTS telephone sets and make telephone calls over the Internet. PC users have a choice of applications that permit them a rich user skill and address book facility, and VoIP telephones are on hand both as desktop models and cordless handsets using Wi-Fi. Mobile roaming users may use their VoIP accounts anywhere they get a broadband Internet connection. As is usually the case in software and systems development, reasonable concentration has not been received by the VoIP security while the development phases and is fall behind in the deployment [2].

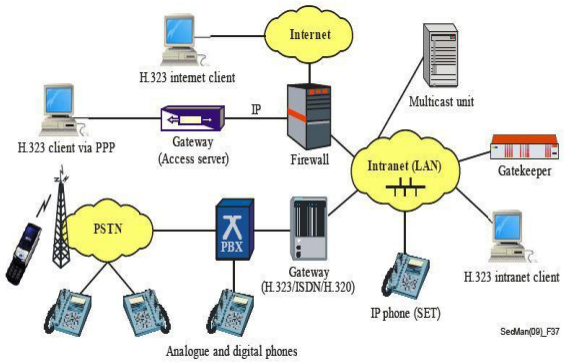


Figure 36 – H.323 system: components and deployment scenarios

3G Technology

Currently there are mostly different views all over the wireless industry as to what constitutes a 3G wireless access network. The problem is swiftly getting worse with the increased usage of 4G to describe, in many cases, technologies that are mainly just evolutions of 3G technologies. Wireless access standards, similar to most other technical standards, usually develop during their service life to put forward enhanced performance and capabilities. The common thought behind different technology "generations" is that each new generation offers important "revolutions" in performance and capabilities compared to its previous technologies. This means that a new

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

"overlay" network, probably in a new frequency band, is required for each technology generation.

In the beginning, Cellular mobile services were offered using analog radio technologies and these were named as the first generation systems called 1G. The designation of 2G was above board because analogue radio networks were put in place of digital ones (2G networks) in the 1990's. However the designation of 3G is not so easy because these various 2G networks have been extensively implemented all over the world and have evolved significantly throughout their long service life to offer greatly enhanced performance and capabilities, mainly for data services.

A. *Function of the International Telecommunication Union In The Designation Of 3G Mobile Standards*

The ITU started work to define the next "generation" of mobile radio standards to shift these networks from National and Regional standards onto a global basis in the mid 1980's. This necessitated discovering a new globally on hand frequency band as well as trying to increase convergence within the several existing 2G wireless technologies. 230 MHz of new radio spectrum was recognized for 'Future Public Land Mobile Telecommunication Systems" (FPLMTS) At the 1992 ITU World Radio Conference, later to be known as International Mobile Telecommunications-2000 (IMT-2000).

Because of the wide deployment and investment in 2G radio technologies during the 1990's IMT-2000 became a "family of standards" offering evolution/revolution options from the main existing 2G network standards. In general an "evolution" opportunity enabled backwards compatible development of a 2G standard to its 3G equivalent within an operators existing spectrum allocation. Whereas a "revolution" option normally required an operator to get extra spectrum, build an overlay network, and utilize dual mode/band mobile equipment.

These 3G ITU standards were finalized in time for 3G services to be firstly launched in 2000. Not amazingly a development option was the first IMT-2000 technology to be deployed.

B. *Large variety of Industry Views on What Constitutes A 3G TECHNOLOGY*

So as to separate 3G from 2G the I*nternational Telecommunication Union* "raised the bar" and defined performance levels appreciably in surplus of those presently obtainable from 2G mobile networks, in particular least data speeds, for a range of specific radio operating environments, were defined. IMT-2000 standards are based on industry submissions which met these new ITU superior performance requisite capabilities. Few of the new "IMT-2000" radio spectrum, recognized in 1992, was auctioned in many countries in the late 1990's for huge sums of money and several country-specific regulations controlled which IMT-2000 family choice could be deployed in these new mobile frequency bands. This naturally resulted in major media focus at the IMT-2000 "revolutionary" family members of standards, which led to rely on several circles that this was the just real 3G.

Actually the "evolutionary" members of the IMT-2000 family enact the huge majority of 3G users at present and are likely to do so for a considerable period of time. This is not at all new in view of the ease of developing to 3G in an operator's existing frequency band, specifically when the 3G technology is backwards compatible with the existing 2G technology, i.e. the 3G network can provide both 2G and 3G users in the same frequency band.

A lot of industry organizations just consider part of the IMT-2000 family of 3G standards as actual 3G technologies in particular IMT-SC (EDGE) is excluded from most 3G mobile statistics. This is mainly fateful because IMT-SC is the "evolutionary" option for the vast installed GSM (2G) base and therefore will almost certainly become the main 3G part in the near future. IMT-SC is usually excluded because many within the industry view CDMA as the only 3G wireless technologies.

C. *IMT-2000 "Evolutionary" 3G Standards*

There are basically two broadly deployed "evolutionary" IMT-2000 standards:

–for evolution from 2G TDMA standards (GSM/IS-136) – IMT-SC (EDGE)

– for evolution from the 2G CDMA Standard IS-95 (cdmaOne) –IMT-MC (cdma2000)

Note that IS-136 can also develop to IMT-MC since it has the similar core network (IS-41).

D. *IMT-2000 "Revolutionary" 3G standards*

These are IMT-2000 standards that normally need operators to get a new spectrum allocation, for example IMT-DS (W-CDMA) because of the relatively large channels (5 MHz), and IMT-TC (TD-SCDMA/UTRA TDD) and IMT-FT (DECT) due to necessity of TDD frequency assignment. Note that it can in several cases be possible to implement IMT-DS in existing cellular bands if enough extra bandwidth can be made available.

E. *Aftermath of Technological Advances*

Early work on 3G in the ITU was directed towards getting a universal spectrum allocation since multi-band radios were at that time economically unattractive. Likewise a single global standard for 3G seemed at the time the only practical solution. Yet it became swiftly clear that even the 230 MHz of new spectrum identified for IMT-2000 in 1992 would be inadequate for future mobile needs.

Because of the fast expansion of 2G mobile during the 1990's it became essential for the ITU to offer a number of possible routes from the different existing 2G systems to a 3G

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

capability. Luckily it also became economically realistic to offer multimode/multiband mobile equipment to smooth the transition from 2G to 3G operations.

IMT-2000 3G wireless technologies definitely have important future development potential, much as 2G technologies have already done, and it seems only reasonable to allow these 3G technologies to fully develop before phasing in a fourth mobile generation.

II. VOIP PROTOCOLS

The two most commonly and widely used network protocols for VoIP are the ITU standard H.323 and the IETF defined SIP. Both are signalling protocols that set up, modify and terminate a VoIP call either unicast or multicast sessions. The Media Gateway Control Protocol (MGCP) provides a signalling and voice control protocol between VoIP gateways and Public Switched Telephone Network (PSTN) gateways. It uses SDP protocol to transmit multimedia streams during a call sessions and RTP (Real Time Transport Protocol).

A. ITU-T H.323 Protocol

H.323 is a standard based on the ITU-T specifications for transmitting calls, video, multimedia transport and data across a network for unicast and multicast conferences. The H.323 standard specification is a protocol suite which includes many sub-protocols [3][7]:

- H.225 for specifying voice controls
- H.235 for providing the security within H.323 and the call setup
- H.245 for control and media stream negotiations.
- H.246 for interoperable support for circuit-switched frameworks.
- H.450 for describing supplementary services such as, call transfer, call on hold and call waiting.

H.235 also addresses security and encryption such as authentication using several algorithms like Diffie-Hellman algorithm, privacy and integrity. It also interoperates with different *H.323* protocols such as H.245 and H.225.
H.323 has four main network elements:

- Terminals: These are the fundamental components of any H.323 architecture. These are endpoints for clients which gives two way communication channels. Every H.323 terminal uses RAS, RTP, H.245 and Q.931 for interacting with the different communication channels and call setup. A terminal can communicate with any other H.323 channel, MCU or any H.323 gateway [8].

- Gateways: A Gateway provides two-way communication between terminals on the Internet Protocol (IP) network and ITU terminals. Gateway is a combination of MGC (Media Gateway Controller) and MG (Media Gateway). MGC manages call signalling and non-media features. MG manages media related functions. A gateway provides H.323 an interface between H.323 and PSTN or other proxy H.323 networks etc [8].
- Gatekeeper: Gatekeeper is very important element of the H.323 system which works like a manager for all calls by acting as a central point. It is used for Call signaling, admission control, address resolution, call authorization, bandwidth management, and ongoing call management [8].
- Multipoint Control Units (MCU): MCU is an endpoint which is responsible for manage multipoint conferences between gateways and terminals. MCU contains mandatory Multipoint controller (MC) and optional Multipoint Processors (MPs). MC handles call signaling and uses H.245 to determine the basic capabilities and functions of the H.323 terminals.

A call establishment is secured and managed by Transport Layer Security (TLS). Once initiated, a call control is established to manage media channel information and encryption. Gatekeeper handles the registered endpoints and permits to place a call. Then, gatekeeper sends the reply by Admission Confirm (ACF) attached with IP address to the calling point. H.323 uses RTP as a TCP over the UDP. Encryption is done inside the packets of RTP through third party. There can be symmetric encryption-based or subscription-based authentication in H.323. In symmetric encryption-based authentication, H.323 protocol applies Diffie-Hellman key-exchange to produce a shared secret ID between two connections or entities. So, prior information and establishment is not required between two communicating devices. But, for subscription-based authentication, shared secret ID is require before the contact between the communicating devices.

Session Initiation Protocol (SIP)

SIP is an application layer protocol which is commonly used to control communication sessions for voice and video calls on Internet Protocol (IP). This protocol is used for establishing call, modifying and terminating calls between unicast or multicast sessions. The architecture of SIP is quite similar to client-server protocol of HTTP thus uses request-response transaction model. Requests are initiated by the client and sent to the server. Server responds the requests and then sends back to the client. SIP relies on the Session Description Protocol (SDP) to carry out the negotiation for codec ID. SIP protocol depends on itself to provide the reliability unlike depending on TCP. It is a text-based protocol like HTTP and SMTP. The SIP system consists of two elements:

- User Agents: A user agent is a logical end-point which is used to send or receive SIP messages.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

This works on behalf of an end-user. SIP User agent can perform the job as a User Agent Client (UAC) which sends requests and another is User Agent Server (UAS) which receive the requests and respond back. This role exists till the duration of the SIP transaction.

- Network Servers: SIP system has a vital component which is a network server. Network servers are of three types. A Proxy server acts as a client (UAC) and a server (UAS) to making request and receiving requests by sending them to the next-hop server. A Registration server is used to receive latest updates on the current locations of the users. It takes REGISTER requests and puts the requests to get the domain IP addresses through SIP URI. A Redirect server is at the receiving requests; it returns the address of the next server or URIs to client rather forwarding the request further.

B. Media gateway Control Protocol (MGCP)

Media Gateway Control Protocol (MGCP) is a call control and signaling protocol which defines the communication between media gateways and Public switched telephonic network (PSTN). This protocol uses RTP for framing the media data and SDP for defining and managing the media streams to transmit into the call sessions. It instructs and allows central coordinator to track the events in IP phones and to send media streams to destination addresses. There is call control intelligence outside the gateways handled by external call agents and then they synchronize between each other for sending accumulated commands to the gateways. It acts as a master-slave protocol. MGCP tries to bring reliability and simplicity and eases for the service providers to design cheap and reliable product.

III. SECURITY ASPECTS OF VOIP

VoIP technology is nowadays widely accepted communication technology. VoIP runs on the internet, so it is quite obvious to inherit the internet security threats. There are possibilities like when communication data of VoIP which is converted into IP packets go through several network connections and access points. So, travelling data can be hacked by the any third party or intruders. There can be many different security threats attached with internet protocols like masqueraders, eavesdroppers, intruders, viruses etc which could be really harmful for the VoIP data.

PC/Laptop based IP phones are more vulnerable to attacks because of specific attack techniques pinpointed to PCs. There can be viruses, malwares, worms, OS vulnerabilities, software applications vulnerabilities etc [4]. Internet Protocol addresses and TCP ports knowledge in attached with packets because voice communication protocols also act like session control protocols. When a NAT technique is used in any network, it becomes difficult to encrypt IP addresses and TCP/UDP port information attached with packets due to information required by NAT for translation. It creates another security breach for these protocols. In H.323 protocol, TCP port 1300 is used to initiate the call connection. But, there is no proper security mechanism applied to secure the establishment of connection. So, this could be dangerous to this protocol. SIP is less vulnerable as they use S/MIME standard to encrypt the establishment of the call connection.

VoIP provides facilitate supplementary services like call forwarding, call divert, park, pick-up, call on hold, conferencing, multi-line etc. Where there is a vulnerability to attack the voice traffic caused by Denial of Service (DoS).

H.323 protocol is still considered and widely implemented by the many manufacturers for voice calls and video conferencing. It is widely used for consumers, business, service providers, entertainment and applications. H.323 standard is designed with four important elements for communication:

Gateways, terminals, multipoint control units and gatekeepers [5]. The networks would be distributed all over the world with the help of their elements. So, there is a possibility on the security aspects of the H.323 as mentioned in the below figure.

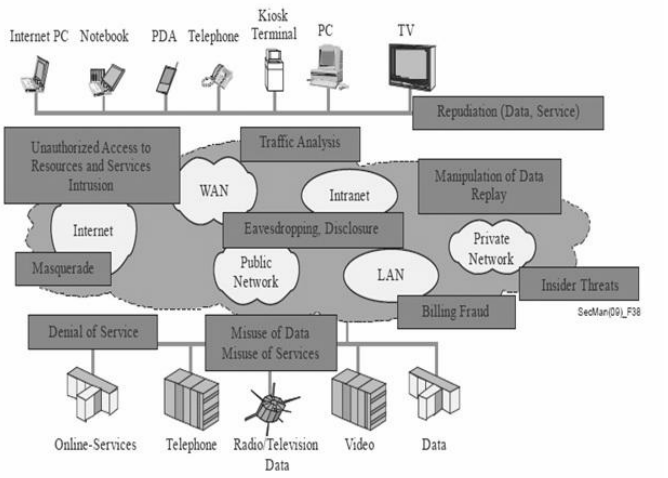


Figure 2: Security Aspects of VoIP

A. The main security aspects in VoIP telephony are as follows:

1. *Server authentication:* Since VoIP users typically communicate with each other using some VoIP infrastructure that involves servers (gateways, gatekeepers, multicast units,), users require to know if they are talking with the correct server and/or with the correct service provider. This applies to both fixed and mobile users.

2. *User/terminal and server authentication:* This is needed to counter security aspects such as connection hijacking, man-in-the-middle attacks, IP address spoofing and masquerade.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

3. *Call authorization*: This is the decision-making process to decide if the user/terminal is actually permitted to use a service feature or a network resource (bandwidth, QoS, codec, etc.). Most often authorization and authentication functions are used together to make an access control decision. Authorization and authentication and help to thwart attacks like masquerade, manipulation, misuse and fraud, and denial-of-service.

4. *Signaling security protection:* This addresses protection of the signaling protocols against manipulation, misuse, confidentiality and privacy. Signaling protocols are typically protected by using encryption as well as by integrity and replay protection measures. Special care has to be taken to meet the critical performance requirements of real-time communication to avoid any service impairment due to security processing.

5. *Key Management:* This includes not only all tasks that are necessary for securely distributing keying material to users and servers, but also tasks like updating expired keys and replacing lost keys. Key management may be a separate task from the VoIP application (password provisioning) or may be integrated with *signaling* when security profiles with security capabilities are being dynamically negotiated and session-based keys are to be distributed.

6. *Inter-domain Security*: This addresses the problem where systems in heterogeneous environments have implemented different security features because of different security policies, different needs and different security capabilities. As such, there is a need to dynamically negotiate security profiles and security capabilities such as cryptographic algorithms and their parameters. This becomes of particular importance when crossing domain boundaries and when different providers and networks are involved. An necessary security requirement for the inter- domain communication is the ability to traverse firewalls smoothly and to cope with constraints of network address translation (NAT) devices.

B. *Major Security Aspects terms:*

Masquerading: A masquerade attack applies a fake identity to gain unauthorized access to use VoIP services. Masquerading can get into charging fraud, breaching of Integrity and privacy. There can be a different ways for masquerading like
Sometimes a user leaves the session or computer open without logging out, so his colleagues or someone else can act as a masquerade attacker. A vulnerable authentication can also lead into an easy cake for attacker to gain access for the confidential data or can modify or steal important data.
So, the best way to overcome this attack is to have write algorithms to have protection shield.

Eavesdropping: Eavesdropping is a type of attack in which an attacker is able to intercept and read the conversations or messages from the user. They are also able to listen to important telephonic conversations. They can also divulge into getting information about the credit card or SSID details.

Denial of Service: A Denial of Service (DoS) is an attack which causes an unavailability of system or network services to users. There can be loss of network connectivity and different network services. It can send large number of requests to services so that the legitimate user would be unable to access the services. DoS decrease the quality of services to the authorized user. It can lead to services interruptions, excessive service data losses, high response delays etc.

Man in the Middle: An attacker is able to read, delete, modify or insert data into the message being transmitted between the two victims without their knowing. The communication between terminals is intercepted by disrupting the TCP connection of an http transaction.

Call hijacking: Call hijacking is an attack in which the calls are redirected to the unauthorized user or hackers by changing the voicemail IP address into hacker-defined IP address. Afterwards, the call is unable to reach to the authorized user. Then, the hacker can mischievously use it to access the confidential data of the legitimate user.

Call Fraud: This type of attack is specific to VoIP and telephonic calls in which it pretends the call is coming from the legitimate user within the network. It uses the VoIP infrastructure to place these calls.

IV. CONCLUSION

The VoIP technology is one of the most popular and fastest growing telecommunication technologies which reduces communication cost as well as better efficiency with less infrastructure costs. In this paper, we have focused on two major telecommunication systems of VoIP technology. We have also talked about common security attacks over H.323 and SIP protocols which make the VoIP technology vulnerable and realize how much we need security solutions for this fast growing cost-effective business. There can be different approach to control security threats like encrypting the voice data passing through the VoIP network. Even though, it also has some limitations. We can also implement firewalls on the data traffic to control the security attacks. We can have a hybrid solution with two or more different security schemes to resolve this issue. We need to ensure about the limitations of the tools and their compatibility issues in different environments.

References

[1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] K. Elissa, “Title of paper if known,” unpublished.

[5] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

An Approach To QoS-Aware Web Service Composition Using Learning Automata

Ali Mehrpour
Department of Computer Engineering,
Science and Research Branch,
Islamic Azad University,
Tehran, Iran

Mir Ali Seyyedi
Department of Computer Engineering,
Science and Research Branch,
Islamic Azad University,
Tehran, Iran

Shahrbanoo Majlesi
Department of Computer Engineering,
Science and Research Branch,
Islamic Azad University,
Tehran, Iran

***Abstract*— Because of growing number of alternative web services that provide same functionality with different qualities, how to select and composite web services to satisfy user's end-to-end constraints is a decision problem. In this paper we have proposed an approach for web service composition based on quality parameters using learning automata consists of two steps: Step1) Stochastic Learning Automata for local selection and Step2) Distributed Learning Automata for global optimization to create composite web service. We have applied these to kind of Learning Automata as a part of Broker in Web Service Architecture. Experimental evaluations show our approach can be applied in dynamic web environment with an acceptable performance without any limitation on number of QoS parameters.**

Keywords-component; Quality of Service (QoS); Web Service Composition (WSC); Stochastic Learning Automata (SLA); Distributed Learning Automata (DLA); Web Service Architecture

I. INTRODUCTION

In recent years, according to development of Service Oriented Architecture (SOA), web services have received much attention. Because of existing alternative web services that provide same functionality with different qualities, creating composite web service from several service units is a decision problem since composite web service must satisfy user's end-to-end QoS requirements [1] [2] [3] [4] such as availability, security, response time and cost. In other words among alternative web services, web services with best QoS must be selected and then service composition plan must be optimized such a way that it not only satisfy user preferences but it also has the highest possible quality.

Learning automata (LA) is one of the important methods in the field of artificial intelligence called machine learning and is used in environments that are not predictable [5] [6]. As the quality parameters of a web service in dynamic web environment change, the use of LA for solving the mention problem is useful.

In this paper we propose an approach for QoS-aware web service composition using learning automata in two steps. In step1, we select locally web services with high quality using Stochastic Learning Automata and then in step2, we optimize globally the composition plan using a Distributed Learning Automata. Also we show the interaction between SLAs and DLA in the architecture of approach. We have applied these two kinds of Learning Automata as a part of Broker in Web Service Architecture to make it more powerful.

The experimental evaluations show our approach can be applied in dynamic web environment where web services QoS parameters are changing constantly with acceptable performance. Furthermore our approach is not dependent on limited number of QoS Parameters.

This paper is organized as follows. In section 2 in this paper, related work about web service composition are discussed. Section 3 describes two kinds of Learning Automata applied in our approach. In section 4 we proposed our architecture and approach in detail. Section 5 shows the experimental evaluations. Finally in the last section, the characteristics of proposed approach and future works have been concluded.

II. RELATED WORK

Up to now different approaches for web service composition are introduced. Approaches are enabled either by workflow research or Artificial Intelligence (AI) planning [7]. The workflow approaches are mostly used in the situation where the request has already defined the process model. The AI planning approaches is used when requester has no process model but has a set of constraints and preferences. Existing approaches based on QoS and user's preferences have some problems. One of the significant problems of these approaches is sufficing to limited number of QoS parameters [8] [9] [10]. Second, these approaches only use QoS information saved in Service Repository by providers which is not confidential since they are not fair. Furthermore there should be consideration of QoS uncertainty and probability that oblige using AI planning. [11] has introduced UDDIe, an extension to Universal Description, Discovery and Integration(UDDI). UDDIe can co-exist with UDDI and support the notation of "blue pages" to record user defined properties associated with a service and to

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

enable discovery of service based on these. In [5] an approach is proposed to enable efficient service selection in dynamic environment of QoS Characteristics using AI planning. First, a Broker-based architecture for web services in which the role of service broker is separated from Service Quality Analyzer is proposed. Second, service selection is performed using LA and the concept of UDDIe by a three-phase technique. Furthermore a solution that combines global optimization with local selection techniques is proposed in [12] [13] to benefit from the advantages of both techniques. Local selection is very efficient but not enables to satisfy global QoS requirement. On the other hand, global optimization can handle global constraints but has poor performance. Proposed approach in [13] has two steps: First, decomposition of global QoS constraints into local constraints using Mixed Integer Programming (MIP). Second, using distributed local selection to find the best web services that satisfy local constraints. To address mentioned problem in this field we propose an approach for QoS-aware web service composition using learning automata.

III. LEARNING AUTOMATA

An automaton is a machine or control mechanism designed to automatically follow a predetermined sequence of operations or respond to encoded instructions [14]. LA as an important algorithm in the field of artificial intelligence is used in situation that the environment is not predictable. In our approach we use two kinds of Learning Automata: Stochastic Learning Automata (SLA) and Distributed Learning Automata (DLA).

A. Stochastic Learning Automata

A stochastic learning automaton [15] is considered as an abstract object with a finite number of actions. SLA selects one of its actions and acts in environment. This action is evaluated by environment and SLA according to the answer, selects the next action. During this process SLA learns to select optimal action. As shown in Fig. 1, LA consists of two parts: 1) SLA with limited number of actions, dealing with a stochastic environment. 2) Learning algorithm by which LA selects its optimal action.

$SLA = \{\alpha, \beta, p, T\}$:

- $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ is a set of actions
- $\beta = \{\beta_1, \beta_2, \dots, \beta_r\}$ is a set of responses (or inputs from environment)
- $P = \{p_1, p_2, \dots, p_r\}$ is the probability vector of actions
- $T = p(n+1) = T[\alpha(n), \beta(n), p(n)]$ is the learning algorithm

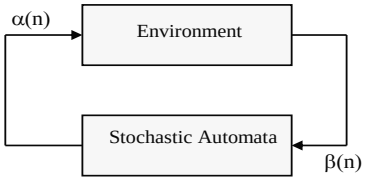


Figure 1. The automata and envrionment

B. Distributed Learning Automata

Distributed Learning Automata (DLA) [16] [17] is a network of LAs that collaborate with each other to solve a specific problem [18]. A DLA with n learning automata is defined as (A, E, T, A_0), where

- $A = \{A_1, A_2, \dots, A_n\}$ is a set of learning automata
- $E \subset A \times A$ is a set of edges in graph in which an edge (i,j) corresponds to action α_j of automaton A_i
- T is a set of learning algorithms
- A_0 is the root automata of DLA

IV. PROPOSED APPROACH

A. Suppositions and Definitions

We suppose there is a Service Repository (SR) which consists of set of Abstract Service Classes. Each Abstract Service Class is allocated to a special functionality and consists of set of web services that realize this functionality.

$$S_j = \{s_1, s_2, \dots, s_m\}, S_j \in SR$$

We also suppose the information related to Service Classes is managed dynamically using UDDIe [8] [11]. Furthermore selected services for composition plan are interoperable with each other.

Definition 1: Abstract Composite Service (ACS) is an abstract representation of a web service composition request which shows set of service classes without pointing to specially web service.

Definition 2: Concrete Composite Service (CCS) is an instance of Abstract Composite Service in which a real web service binds to each web service class.

Definition 3: For expressing quality of each web service, we use QoS Vector (Q_{S_j}).

$$Q_{S_j} = \{q_1, q_2, \dots, q_r\}, \quad q_i, \quad 1 \le i \le r$$

Where value of q_i is quality value of web service s_j for q_i attribute. These values can be obtained either directly from service provider such as cost or dynamically using historical executions or user's feedbacks such as response time.

Definition 4: QoS Global Constraints (GC) are QoS values of composite service which are specified by service requester. CCS must be satisfied this GC.

$$GC = \{c_1, c_2, \dots, c_k\}$$

Definition 5: QoS Local Constraints (LC). For local selection, each global constraint ($c_i, 1 \le i \le k$) is decomposed to n local constraints where n is the number of service classes.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Definition 6: Web Service Ranked List (RL). In local selection services must be selected that satisfy local constraints. Web service ranked list with functionality of j is showed by S_j.

$$RL\,S_j = \{s_1, s_2, \dots, s_w\}$$

In this ranked list, according to satisfying all local constraints, s_1 has maximum profitability and s_w has minimum profitability in comparison with other web services.

B. Proposed Architecture

Proposed architecture is showed in Fig. 2. In this architecture, the concept of learning automata is used in two levels to make broker more powerful; in order to responds to user's request better with more performance. In first level we used k number of SLAs which receive a service class with local constraints. The number of SLAs is equal with number of service classes in ACS. Each SLA is delimited to definite service class and its duty is supplying a Ranked List of web services with definite functionality. In second level we used a DLA which receive Ranked List from first level and then optimizes composition plan.

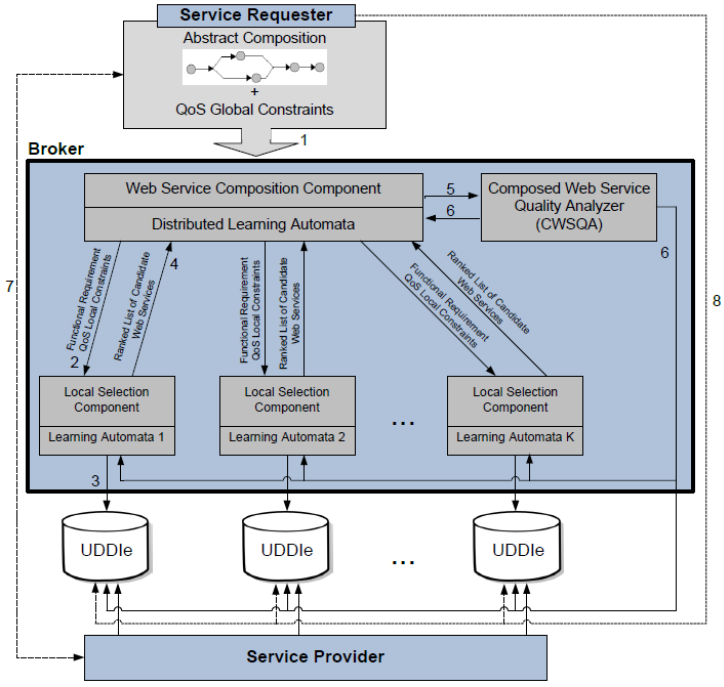


Figure 2. Our proposed architecture

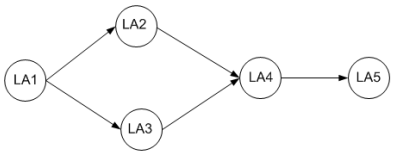


Figure 3. DLA of a composite service

C. Mapping Problem to Learning Automata

To solve QoS-Aware Web Service Composition problem we use a DLA besides k number of SLA where k is changeable according to the number of service classes requested in ACS. These LAs Cooperates each other via messages. Actions of each LA are Ranked List of web services with definite functionality that definite SLA in first level creates. Furthermore Composite Web Service Quality Analyzer (CWSQA) is the environment of DLA. DLA according Fig. 3 act as:

First, LA_1 selects one of its actions based on QoS probability vector and sends messages to LA_2, LA_3 and makes them active. Rest SLAs act as same as LA_1. Finally LA_5 actives, LA_5 selects one of its actions and then created composite web service is delivered to the CWSQA. After QoS calculation of composite service, response is delivered to DLA by environment. Then LAs in DLA update the probability vector of their actions by learning algorithm. The learning algorithm of LAs of DLA is Reward-Penalty showed in relation 1.

$$If\ \alpha(n) = \alpha_i, \qquad (1)$$

$$when\ \beta = 0 \quad \begin{cases} p_j(n+1) = (1-a).p_j(n) & for\ all\ j \neq i \\ p_i(n+1) = p_i(n) + a.[1 - p_i(n)] \end{cases}$$

$$when\ \beta = 1 \quad \begin{cases} p_j(n+1) = \dfrac{b}{r-1} + (1-b).p_j(n) & for\ all\ j \neq i \\ p_i(n+1) = (1-b).p_i(n) \end{cases}$$

SLAs in second level act as same as learning automata proposed in [5] with a difference; these SLA only calculate the utility of web services which satisfy local constraints by using Utility Function, So ineffective web services are not considered in global optimization. In each cycle, each SLA calculate utility of web services which is able to satisfy local constraints according to QoS value existed in service repository and then creates a ranked list of web services based on these utilities, then each web service is compared with its last position in ranked list. If the new position of web service improves it will be reward else it will be penalty. Learning algorithm of these SLAs is L_{r-p} showed in relation 1.

In order to doing local selection and then global optimization, local constraints must be decomposed to local constraints. Decomposition should be done effectively in order to consider all useful candidate web services for global optimization. To decompose, we use method proposed in [19].

To create web service ranked lists in first level, we need a Utility Function. Utility Function specifies which web service is more effective according its quality parameters. Utility Function is dependent to the concept of Multiple Attribute Decision Making and the goal of it is utility calculation of a web service according to QoS parameters which have different measures and values. There are several Utility Functions [19]. A simple method that we use is Simple Additive Weighting (SAW).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

D. Proposed Algorithm

Input:

AWS: Abstract web service
GCs: Global constraints
S: Set of candidate services for each service class with service information in UDDIe

Output:

CCS: Concrete composite service

Algorithm:

1. Decompose GCs to local constraints and are sent to first level SLAs.
2. First level SLAs based on past learns, deliver Ranked Lists of web services to SLA according local constraints and service information in UDDIe.
3. DLA uses ranked lists and do learning until reaching to desirable responses to user.

   3.1. SLAs in DLA select web service with maximum probability and new composite service is generated and evaluated by WSQA.

   3.2. According to environment (CWSQA) responses, each service in composite service Reward or Penalty and also probability vector of each SLA in DLA has been updated.

   3.3. Information of web service in UDDIe has been updated.

V. Experimental Evaluation

We have done the simulation using C# programming language and Microsoft visual studio 2010 IDE. System properties are:

- **CPU:** Intel® Core™ 2 Duo CPU T9300@ 2.5GHz
- **Memory(RAM):** 4.00 GB
- **System Type:** 32-bit Operating System (Windows 7)

In our evaluation we used QWS[1] dataset. QWS comprises measurement of 9 QoS attributes for 2500 real web service. These services were collected from public sources on the web, including UDDI registries, search engines and service portals, and their QoS values were measured using commercial benchmark tools. More details about this dataset can be found in [20].

Although proposed approach is not dependent to limited number of QoS parameters; we have used only one parameter, response time, for simplicity. We have also use sequential composition model.

Fig. 4 illustrates relationship between number of candidate web service and percentage of DLA convergence to response. The results show our approach on average more than 90% converges to response.

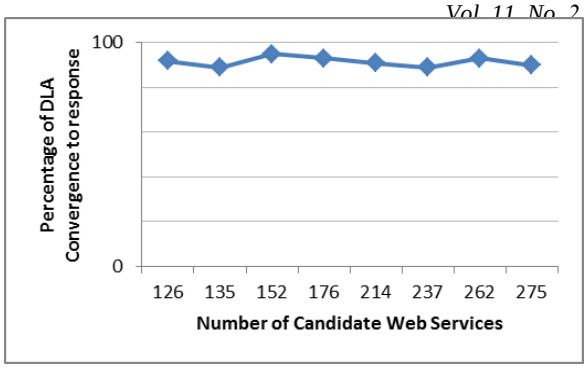


Figure 4. Percentage of DLA convergence to response vs. number of candidate web services

Fig. 5 compares our approach with proposed approach in [12] according to execution time. Also our approach profits from acceptable execution time vs. an approach that optimizes composite plan according all candidates web services, but in comparison with [12] has more execution time because using LAs in order to support dynamic web environment.

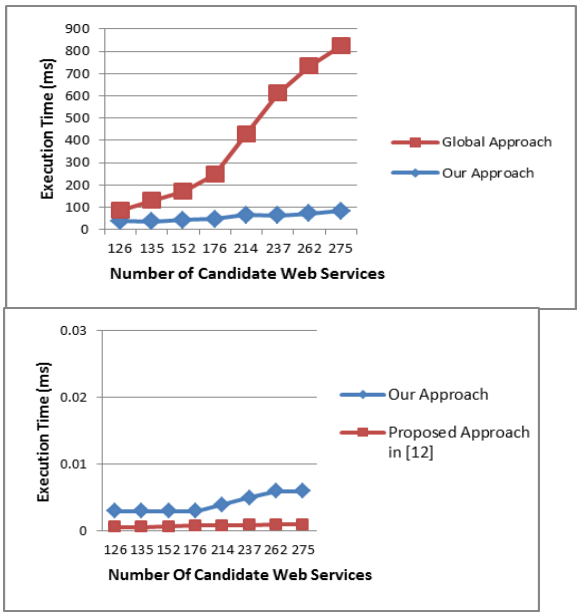


Figure 5. Execution time vs. number of candidate web services

Fig. 6 shows the deviation of DLA responses from best responses. In this Figure we have sampled 71 cases that DLA has converged to response and then we have measured how much the DLA response deviates from best response. The result shows learning process causes DLA finds more appropriate responses in final requests.

[1] Available at http://www.uoguelph.ca/~qmahmoud/qws/index.html

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

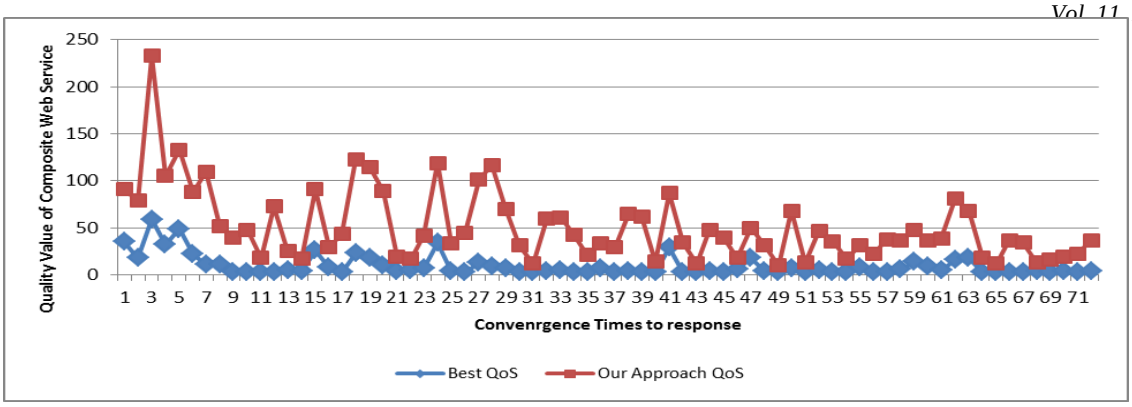


Figure 6. Deviation of our approach QoS from the best QoS

VI. CONCLUSIONS

To satisfy user's end-to-end QoS constraints in web service composition problem, we have proposed an approach using Learning Automata in two levels. SLAs in first level perform local selection and then DLA in second level uses the results of SLAs to perform global optimization. Simulation result showed by using this approach we profit not only from more performance in dynamic web composition but also from service executions history since information provided by service provider is not reliable completely. Furthermore our approach is not dependent on limited number of QoS parameters.

Our future works will be related to do more simulation for complicated composition model in order to do more careful evaluation and improve learning algorithm. Furthermore an effective method for service repository management is necessary to save web service information updating constantly.

VII. REFERENCES

[1] Alrifai, Mohammad and Risse, Thomas. *Efficient QoS-aware Service Composition.* Hanover, Germany : IEEE, 2008.

[2] Papazoglou, et al., et al. *Service-Oriented Computing: A Research Roadmap.* 2008, International Journal Of Cooperative Information Systems, Vol. 17, pp. 223-255.

[3] Liu, Bing, Shi, Yuliang and Wang, Haiyang. s.l. *QoS Oriented Web Service Composition and Optimization in SOA.* IEEE, 2009. Joint Conferences on Pervasive Computing (JCPC). pp. 605-610.

[4] Pejman. E., Rastegari Y., Majlesi Esfahani P. and Salajegheh A. *Web Service Composition Methods: A Survey.* In Proceeding of International MultiConferences of Engineers and Computer Scientists. 2012 VOL I, IMECS 2012, March 14-16, 2012, Hong Kong.

[5] Tabein, Reza, Moghadasi, Mahdi Naser and Khoshkbarforoushha, Alireza. *Broker-based Web Service Selection using Learning Automata.* s.l. : IEEE, 2008. International Conference on Service Systems and Service Management. pp. 1-6.

[6] Zhang, Xiwen, et al. *A Learning Automation Solution to the QoS-Aware Service Composition.* Shanghai : IEEE, 2009. Web Information Systems and Mining, 2009. WISM 2009. International Conference. pp. 297 - 301.

[7] Rao, Jinghai and Su, Xiaomeng. *A Survey of Automated Web Service Composition Methods.* s.l. : Springer-Verlag, 2005, Vol. LNCS 3387, pp. 43-54.

[8] Liu, Yutu, Ngu, Anne H.H. and Zeng, Liangzhao. *QoS Computation and Policing in Dynamic Web Service Selection.* New York : Proceedings of the Thirteenth International World Wide Web Conference, May 2004. pp. 66-73.

[9] Sheth, Amit, et al., et al. *QoS for Service-oriented Middleware.* Orlando : Proceedings of the Conference on Systemics, Cybernetics and Informatics, July 2002.

[10] Zeng, Liangzhao, et al., et al. *QoS-Aware Middleware for Web Services Composition.* s.l. : IEEE Transaction on Software Engineering , 2004. Vol. 30, pp. 311-327.

[11] SheikhAli, Ali, et al., et al. *UDDIe: an extended registry for Web services.* s.l. : IEEE, Applications and the Internet Workshops, 2003. pp. 85-89.

[12] ALRIFAI, MOHAMMAD, RISSE, THOMAS and NEJDL, WOLFGANG. *A Hybrid Approach for Efficient Web Service Composition With End-to-End Constraints.* Hanover, Germany : ACNM Transactions on Web, May 2012. Vols. 6, No 2.

[13] Alrifai, Mohammad and Risse, Thomass. *Combining Global Optimization with Local Selection for Efficient QoS-aware Service Composition.* Madrid : Proceeding of the 18th international conferences on World wide web, 2009. pp. 881-890. 978-1-60558-487-4.

[14] Ünsal, Cem. *Intelligent Navigation of Autonomous Vehicles in an Automated Highway System.* Doctor of Philosophy in Electrical Engineering, 1997, Blacksburg, Virginia.

[15] NARENDRA, KUMPATI S and THATHACHAR, M A. L. *Learning Automata - A Survey.* JULY 1974, Vols. SMC-4, 4, pp. 323-334.

[16] Mølsæther Stensby , Aleksander and Moy, Ole-Alexander. *Distributed Learning Automaton.* MAY 2007, Agder University College.

[17] Beigy, Hamid and Meybodi, Mohammad Reza. *A New Distributed Learning Automata Based Algorithm For Solving Stochastic Shortest Path Problem.* Durham, USA : Proceedings of the Sixth International Joint Conference on Information Science, 2002. pp. 339-343.

[18] Friedman, Eric J. and Shenker Scott. *Learning by Distributed Automata.* Departement of Industrial Engineering and Operations Research University of California, Berkeley, CA 94720. MAY 1993.

[19] Yoon, Paul K. and Hwang, Chinq Lai. Multiple Attribute Decision Making: An Introduction. s.l. : Sage Publications, Inc; illustrated edition edition, 1995.

[20] AL-MASRI, E. and MAHMOUD, Q. H. *Investigating web services on the world wide web.* New York : Proceeding of the 17th international conferences on World wide web, 2008. pp. 795-804.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Demonstration of the Functioning of TCP Protocol Used for Network Congestion Control

Asagba, Prince Oghenekaro
Department of Computer Science
University of Port Harcourt
Port Harcourt, Nigeria

Anucha, Udo Sylvester
Department of Computer Science
University of Port Harcourt
Port Harcourt, Nigeria

Ogini, Nicholas Oluwole
Department of Mathematics and Computer Science
Delta State University
Abraka, Nigeria

Abstract — **Congestion can occur when the quality of service in a network reduces as a result of a node or link conveying too many data. TCP is the most widely used protocol for Internet traffic, including email, web browsing, data and an increasing portion of multimedia content delivered in real time using the HTTP/TCP protocols. Performances of existing TCP congestion control algorithms degrade significantly when deployed over wireless networks. TCP was designed primarily for reliability as opposed to real time delivery, but the problem is particularly severe for real time applications, such as, HTTP/TCP based streaming. In this paper, we carried out a research on the TCP's four related congestion control algorithms, namely: slow-start, congestion avoidance, fast retransmit and fast recovery. We studied the behaviour and implementation of slow-start and congestion avoidance algorithms, as well as the modifications to the fast retransmit and fast recovery. We used the OPNET Network Model as our methodology. The TCP performance on the network was modeled, first without background traffic and then with background traffic. We compared these algorithms using the same network model to deterministically check several scenarios; and simulations were conducted to ascertain the differences between the congestion control algorithms studied and OPNET's software. The results gotten showed that using different algorithms, traffic could actually be fine tuned in the network being modeled so as to achieve higher Performance. The adjustments were done in the OPNET simulator.**

Keywords - TCP Protocols; Congestion control algorithms; Network; Acknowledgment (ACK); OPNET Network

I. INTRODUCTION

Congestion occurs when there are too many sources sending too much data too fast for the network to handle, and it is a serious problem. Congestion control is the efforts made by network nodes to prevent or respond to overload conditions [9]. Congestion can also occur when the quality of service in a network reduces as a result of a node or link conveying too many data.

Congestion control keeps a set of senders from sending too much data into network because of lack of resources at some point. Congestion control and resource allocation involves both host and network elements such as routers, switches, computer systems (clients and servers). TCP is the dominant transport protocol of today. It does not meet demand for fast transfer of large volumes of data and the deployment of the network infrastructures that is ever increasing, because it favours reliability over timeliness and fails to fully utilize the network capacity due to limitations of its conservative congestion control algorithm [4]. Congestion control algorithms are measures in handling traffic from a node or link conveying too many data in a network to effectively manage its carrying capacity.

TCP establishes a full duplex virtual connection between two endpoints. Each endpoint is defined by an IP address and a TCP port number. TCP sends a full window of information at the beginning of the transmissions. In the same way, when a packet is dropped, the destination cannot acknowledge further segments until the lost packet arrives; therefore the source will probably run out of window and will have to wait until a timeout halves the send window and forces a retransmission; after that, a cumulative acknowledgement will be received which will free space on the send window (probably opening it completely), so that a full window can be transmitted together, as in the first case. The algorithms specified in this paper gives notification for congestion whenever there is a loss in the network. This concept is used also by Explicit Congestion Notification (ECN). The assumption is that, for a loss to occur there must have been congestion. We also demonstrated the functioning of TCP protocol, and particularly compared the four algorithms used for congestion control: slow start, congestion avoidance, fast retransmit and fast recovery. We presented a number of scenarios to simulate and compare these algorithms.

The slow start algorithm is used for this purpose at the beginning of a transfer, or after repairing loss detected by the retransmission timer. Slow start additionally serves to start the "ACK clock" used by the TCP sender to release data into the network in the slow start, congestion avoidance, and loss recovery algorithms [3].

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

II. LITERATURE REVIEW

Various studies show that up to 80% of existing Internet multimedia services are HTTP/TCP based [11]. Congestion control algorithm is an integral component of TCP that directly determines the performance of the protocol. Based on the inputs used by the congestion control algorithms, TCP congestion control algorithms can be categorized into slow start, congestion avoidance, fast retransmit and fast recovery algorithms [7]. Each of these is described as thus:

A. Slow-Start

Slow-start algorithm is used when congestion window (cwnd) < slow start threshold (ssthresh), while the congestion avoidance algorithm is used when cwnd > ssthresh. When cwnd and ssthresh are equal, the sender may use either slow start or congestion avoidance [1][2]. The SYN/ACK and the acknowledgment of the SYN/ACK must not increase the size of the cwnd. Furthermore, if the SYN or SYN/ACK is lost, the initial window used by a sender after a correctly transmitted SYN must be one segment consisting of at most Sender Maximum Segment Size (SMSS) bytes [5]. Much of the difficulties in understanding these sizes and their relationship had been that of the variable sizes of the IP and TCP headers. Both protocols have varying sizes.

During slow start, a TCP increments cwnd by at most SMSS bytes for each ACK received that cumulatively acknowledges new data. Slow start ends when cwnd exceeds ssthresh (or, optionally, when it reaches it, as noted above) or when congestion is observed. While traditionally TCP implementations have increased cwnd by precisely SMSS bytes upon receipt of an ACK covering new data, we RECOMMEND that TCP implementations increase cwnd, as shown below:

cwnd += min (N, SMSS)

where N is the number of previously unacknowledged bytes acknowledged in the incoming ACK [15]. This adjustment is part of Appropriate Byte Counting and provides robustness against misbehaving receivers that may attempt to induce a sender to artificially inflate cwnd using a mechanism known as "ACK Division". ACK Division consists of a receiver sending multiple ACKs for a single TCP data segment, each acknowledging only a portion of its data. A TCP that increments cwnd by SMSS for each such ACK will inappropriately inflate the amount of data injected into the network [10]. Hosts are not required to reassemble infinitely large TCP datagrams. So, slow-start is part of the congestion control strategy used by TCP, the data transmission protocol used by many Internet applications. Slow-start is used in conjunction with other algorithms to avoid sending more data than the network is capable of transmitting, that is, to avoid causing network congestion.

B. Congestion Avoidance

The slow start and congestion avoidance algorithms must be used by a TCP sender to control the amount of outstanding data being injected into the network. To implement these algorithms, two variables are added to the TCP per-connection state. The cwnd is a sender-side limit on the amount of data the sender can transmit into the network before receiving an acknowledgment (ACK), while the receiver's advertised window (rwnd) is a receiver-side limit on the amount of outstanding data. The minimum of cwnd and rwnd governs data transmission [14]. Another state variable, the slow start threshold (ssthresh), is used to determine whether the slow start or congestion avoidance algorithm is used to control data transmission [3][5]. The cwnd is not to be confused with the TCP window size which is maintained at the receiver's side**.**

C. *Fast Reco*very

A TCP receiver SHOULD send an immediate duplicate ACK when an out-of-order segment arrives. The purpose of this ACK is to inform the sender that a segment was received out-of-order and which sequence number is expected. From the sender's perspective, duplicate ACKs can be caused by a number of network problems. First, they can be caused by dropped segments. In this case, all segments after the dropped segment will trigger duplicate ACKs until the loss is repaired. Second, duplicate ACKs can be caused by the re-ordering of data segments by the network (not a rare event along some network paths). Finally, duplicate ACKs can be caused by replication of ACK or data segments by the network. In addition, a TCP receiver SHOULD send an immediate ACK when the incoming segment fills in all or part of a gap in the sequence space. This will generate more timely information for a sender recovering from a loss through a retransmission timeout, a fast retransmit, or an advanced loss recovery algorithm [12]. This is a way of stopping the link between the source and destination from getting overloaded with too much traffic.

D. Fast Retransmit

The TCP sender SHOULD use the "fast retransmit" algorithm to detect and repair loss, based on incoming duplicate ACKs. The fast retransmit algorithm uses the arrival of 3 duplicate ACKs as an indication that a segment has been lost. After receiving 3 duplicate ACKs, TCP performs a retransmission of what appears to be the missing segment, without waiting for the retransmission timer to expire. After the fast retransmit algorithm sends what appears to be the missing segment, the "fast recovery" algorithm governs the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

transmission of new data until a non-duplicate ACK arrives. The reason for not performing slow start is that the receipt of the duplicate ACKs not only indicates that a segment has been lost, but also that segments are most likely leaving the network (although a massive segment duplication by the network can invalidate this conclusion). In other words, since the receiver can only generate a duplicate ACK when a segment has arrived, that segment has left the network and is in the receiver's buffer, so we know it is no longer consuming network resources. Furthermore, since the ACK "clock" is preserved, the TCP sender can continue to transmit new segments (although transmission must continue using a reduced cwnd, since loss is an indication of congestion) [6][8]. By comparing its own congestion window with the received window of the receiver, a sender can determine how much data it may send at any giving time.

III. METHODOLOGY

We began by comparing two algorithms slow-start and congestion avoidance. We used the same network model to deterministically check several scenarios for the comparison of the four algorithms. Figure 1 shows an OPNET Network Model for the study. It is a deterministic network model used to try several IF's scenarios for the comparisons. The model comprises of a Profile Definition, an Application Definition, five client systems in a switched LAN, five server machines in another switched LAN network, both networks linked through their gateways (routers) to the cloud (Internet). The two LANs can be separated over a geographic location.

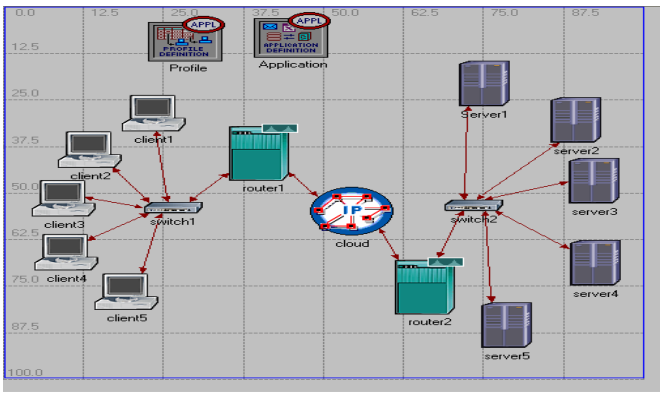


Fig. 1: OPNET Network Model for the study

Previous versions of TCP start a connection with the sender injecting multiple segments into the network, up to the windows size advertised by the receiver [16]. When the hosts are placed on the same LAN, the result may be okay, but the slow-start algorithm could be used as a remedy if intermediate slower connections and routers are placed between source and destination, the data / packets must queue because there is a possibility for the links to run out of storage in the queue.

Beginning transmission into a network with unknown conditions requires TCP to slowly probe the network to determine the available capacity, in order to avoid congesting the network with an inappropriate large burst of data [4]. Figure 2 gives the simulation sequence for the slow-start and congestion avoidance algorithms, showing the time average of file transfer protocol (FTP) and the download response time captured over a period of 30 simulated minutes.

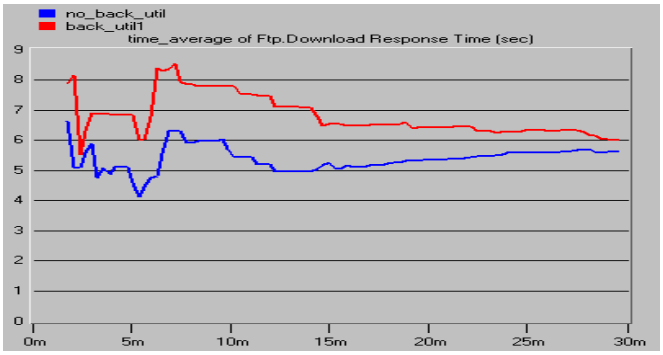


Fig. 2: Download response time (sec)

Slow start adds another window to the sender's TCP, the congestion window (cwnd). When a new connection is established with a host on another network, the cwnd is initialized to one segment (typically 536 bytes or 512 bytes). The sender starts by transmitting one segment and waiting for its ACK. When that ACK is received, the cwnd is increased from one to two, and two segments can be sent. When each of these two segments is acknowledged, the congestion window is increased to four. This provides an exponential growth, although it is not exactly exponential because the receiver may delay its ACK's, typically sending one ACK every two segments that it receives. The sender can transmit up to the minimum of the congestion window and the advertised window [10]. The congestion window is flow control imposed by the sender, while the advertised window is flow control imposed by the receiver. At some point, the capacity of the Internet can be reached and an intermediate router will start discarding packets. This tells the sender that its congestion window has gotten too large.

Congestion avoidance is a way to deal with lost packets. Congestion can occur when data arrives on a big pipe (a fast LAN) and outputs on a smaller pipe (a slower WAN). Congestion can also occur when multiple input streams arrive at a router whose output capacity is less than the sum of the inputs. There are two indications of packet loss at a sender: a timeout occurring and the receipt of duplicate ACK's. However, the overall assumption of the algorithm is that packet loss caused by damage is very small (much less than

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

1%); therefore the loss of a packet signals congestion somewhere in the network between the source and destination [13]. Although congestion avoidance and slow start are independent algorithms with different objectives, in practice, they are implemented together. When congestion occurs TCP must slow down its transmission rate of packets into the network, and then invokes slow start to get things going again [3]. This means that if all segments are received and the acknowledgments reached the sender on time, everything must have worked well.

The combined congestion avoidance and slow start algorithms require that two variables are maintained for each connection: a congestion window (cwnd) and a slow start threshold size (ssthresh).

The combined algorithm operates as follows [4]:

1) For author/s of only one affiliation Initialization for a given connection sets cwnd to one segment and ssthresh to 65535 bytes. The initial value of cwnd must be less than or equal to 2*SMSS bytes and must not be more than 2 segments. SMSS is the size of the largest segment that the sender can transmit. The initial value of cwnd may be arbitrarily high (some implementations use the size of the advertised window), but it may be reduced in response to congestion.

2) The TCP output routine never sends more than the minimum of cwnd and receiver's advertised window.

3) When congestion occurs one-half of the current window size is saved in ssthresh. Additionally, if the congestion is indicated by a timeout, cwnd is set to one segment. Congestion is indicated by a timeout or the reception of duplicate ACK's.

4) When new data is acknowledged by the other end, it increases cwnd. The way in which cwnd is increased depends on whether TCP is performing slow start or congestion avoidance. If cwnd is less than or equal to ssthresh, TCP is in slow start, otherwise TCP is performing congestion avoidance.

Slow start continues until TCP is halfway to where it was when congestion occurred, and then congestion avoidance takes over. This is done due to the recorded half of the window size that caused the problem.

From the foregoing, slow start increases congestion cwnd exponentially. Congestion avoidance on the other hand dictates that cwnd be incremented by segsize * segsize / cwnd each time an ACK is received, where segsize is the segment size and cwnd is maintained in bytes. This results in a linear growth of cwnd, compared to slow start's exponential growth. The increase in cwnd should be at most one segment each round-trip time (RTT), regardless how many ACK's are received in that RTT whereas slow start increments cwnd by the number of ACK's received in a RTT [2].

This network setup, utilizes TCP as its End-to-End transmission protocol. Five servers are placed in one side of geographical location and five clients are placed in the other side. The Congestion window size will be analyzed with different mechanism. This network is assumed to be perfect with no packet loss. Figure 3 shows congestion avoidance versus slow start, the time average of point to point utilization as experienced in Slow Start and Congestion Avoidance.

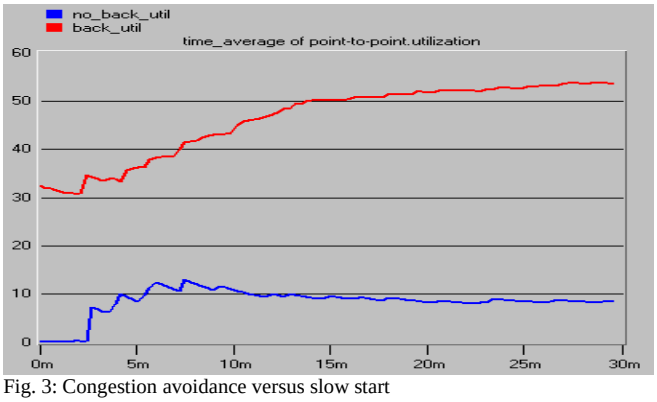


Fig. 3: Congestion avoidance versus slow start

Figure 4 shows the simulated control of congestion, result gotten when the two scenarios were duplicated. We observed that the first scenario (red) exhibited high congestion, which was taken care of at the second scenario when the parameters where adjusted to control congestion (congestion avoidance).

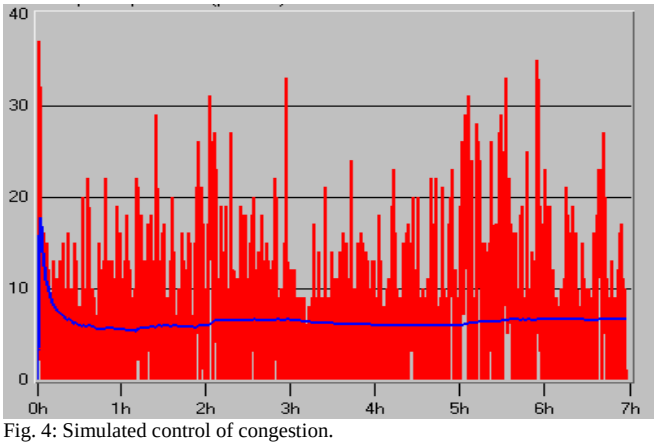


Fig. 4: Simulated control of congestion.

IV. RESULTS

From our analysis, we found out that:

The congestion control algorithms studied can be interpreted in slightly different ways. These interpretations give name to several TCP flavours such as Tahoe, Reno, and New-Reno. Simulations were conducted in order to clarify what the differences are between the congestion control

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

algorithms studied and OPNET's software. The OPNET's software has shown itself as a perfect tool for achieving such goal. Some remarks can be made after analyzing simulation's results.

- Tahoe TCP provides us better performance than Reno TCP, because the former closes the usable window when the first error of a burst is detected, and uses Slow Start from the beginning.
- However, New Reno TCP overcomes both algorithms since it avoids closing the usable window when more than one error occurs. Although it may still requires further studies. These simulations show that the best TCP flavour will be New Reno TCP using selective acknowledgments. Figure 5 shows the overall channel throughput as result of all the adjustments made on all algorithms, vis-à-vis congestion control.

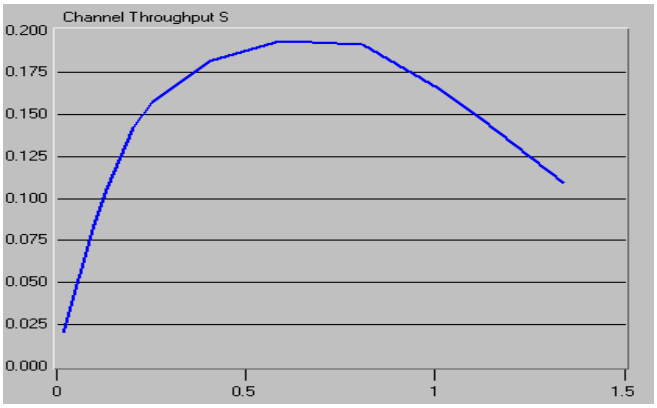


Figure 5: Overall channel throughput

V. CONCLUSION

The objectives of the research, was to model a WAN composed by two LANs, and determine how the background traffic is affecting TCP traffic on the network. Hence the TCP performance on the network was modeled, first without background traffic and then with background traffic. Because there is no interest in modeling the details of each LAN we used available LAN models to model the individual LANs as five nodes each. The results gotten showed that using different algorithms, traffic can actually be fine tuned (small adjustments can be made on the traffic using different algorithms for optimal performance and effectiveness) in the network being modeled so as to achieve the overall best results.

REFERENCE

[1] M. Amirijoo, P. Frenger, F. Gunnarsson, M. Johan and Z. Kristina, "On self-optimization of the random access procedure in 3g long term evolution," Ericsson Research, Ericsson AB, Sweden, pp. 177-184, 2009.

[2] R. Boder and C. G. Lee, "Real-time guarantee of a periodic packets in single-hope ad hoc wireless networks," IEEE Int'l Conference on Embedded and Real-time Computing Systems and Applications, pp. 254-259, 2005.

[3] Y. Choi, "Multichannel random access in ofdma wireless networks," IEEE Selected Areas in Communications, vol. 24, pp. 603-613, 2006.

[4] D. Clark, "Windows and acknowledgement strategy in TCP, ARPANET working group requests for comment," DDN Network Information Center, SRI International, Menlo Park, CA. RFC-813, pp. 8-19, 1982.

[5] D. E. Comer, " Internetworking with TCP/IP: principles, protocols, and architecture," 1 (5th ed.), Prentice Hall, ISBN 0-13-187671-6, pp. 98-101, 2006.

[6] W. B. Dunbar, "A distributed receding horizon control algorithm for dynamically coupled nonlinear systems," IEEE Conference on Decision and Control, pp. 3-4, 2005 .

[7] S. W. Edge, "An adaptive timeout algorithm for retransmission across a packet switching network," Proceedings of SIGCOMM '83 (Mar. 1983), ACM, pp. 174-179, 1983.

[8] G. Hauksson and M. Alanyali, "Wireless medium access via adaptive backoff: delay and loss minimization," IEEE 27th Conference on Computer Communications, pp. 1777-1785, 2008.

[9] C. Langbort, R. S, Chandra and R. D'Andrea, "Distributed control design for systems interconnected over an arbitrary graph," IEEE Trans. on Automatic Control, 49(9):1502-1519, pp. 9-12, 2004.

[10] Q. Ling and M. D. Lemmon, "Robust performance of soft real-time networked control systems with data dropouts," IEEE Conference on Decision and Control, pp. 1225-1230, 2002.

[11] J. Nagle, "Congestion control in IP/TCP internetworks ARPANET working group requests for comment," DDN Network Information Center, SRI International, Menlo Park, CA, Jan. 1984, RFC-896. pp5, 1984.

[12] T. P. Ruggaber and J. W. Talley, "Detection and Control of Combined Sewer Overflow Events using Embedded Sensor Network Technology," Proceedings of the World Water and Environmental Resources Congress, pp. 101-110, 2005.

[13] G. Sharma and A. Ganesh, "Performance analysis of contention based medium access control protocols," IEEE Trans. Inf. Theory, vol. 55, pp. 1665 -1682, 2009.

[14] Y. Sun and M. D. Lemmon, "Periodic communication logics for the decentralized control of multi-agent systems," IEEE Conference on Control Applications, pp. 1431-1434, 2005.

[15] M. Velasco, J. M. Fuertes and P. Marti, "The self triggered task model for real-time control systems," IEEE Work-in-Progress Session of the 24th Real-Time Systems Symposium (RTSS03), pp. 2-3, 2003.

[16] Y. Yang and R. Kravets, "Distributed QOS guarantees for real-time traffic in ad hoc networks," IEEE Conference on Sensor and Ad Hoc Communications and Networks, pp. 118-127, 2004

AUTHORS PROFILE

Prince Oghenekaro Asagba had his B.Sc. degree in Computer Science at the University of Nigeria, Nsukka, in 1991, M.Sc. degree in Computer Science at the University of Benin in April, 1998, and a Ph.D degree in Computer Science at the University of Port Harcourt in March, 2009. He is a Senior Lecturer and a visiting lecturer to several Universities in Nigeria since 2010. His research interest include: Network Security, Information Security, Network Analysis, Modeling, Database Management Systems, Object-oriented Design, and Programming. He is a member of Nigeria Computer Society (NCS) and Computer Professional Registration Council of Nigeria (CPN).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Udo Sylvester Anucha obtained his first degree - Bachelor of Engineering (B.Eng) in Computer Engineering from Enugu State University of Science and Technology, M.Sc. degree in Computer Science at the University of Port Harcourt I 2011. He is a PhD student in the University of Port Harcourt, Port Harcourt, Rivers State, Nigeria. He holds some IT professional certifications which include: MCSE, MCITP, MCP, MCTS, CCNA and A+. He is actively involved in researches on throughput performance of wireless networks.

Nicholas Oluwole Ogini received his B.Sc (1993), M.Sc (1998), and Ph.D (2013) in Computer Science from the University of Benin. He is currently a lecturer at the Delta State University, Abraka, Nigeria as lecturer I. His research interests include: Information Security, Database Management Systems, Fuzzy Expert systems, programming. He is a member of Nigerian Computer Society (NCS) and Computer Professional Registration Council of Nigeria (CPN).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Change management Strategies and Processes for the successful ERP System Implementation: A Proposed Model

Abdullah Saad AL-Malaise AL-Ghamdi
Department of Information Systems
Faculty of Computing and Information Technology
King Abdul Aziz University

***Abstract*— Recent advancement in information technology and business development, the business organizations turned towards the adoption of advanced information technology systems for their organizational setup. Progression of technologies in business environment has been observed in many organizations by the initiation of enterprise resource planning (ERP) system implementation. ERP is business integrated information system software that attracts the attention of business organizations in order to improve their business processes and achieve the company's goals. Almost all the ERP system implementation is based on change management system, where the traditional/ legacy system is completely replaced with the new and advance system. This paper will discuss the change management strategies and processes for the success of ERP system implementation. The paper has proposed a model, change management strategies and processes for the successful ERP system implementation that will strengthen the scope of the title of this paper.**

Keywords-component; Change Management, IT, ERP, User Reaction, System, Implementation Process

I. INTRODUCTION

Change management system is one of the most common and critical success factors of ERP implementation (Aladwani, 2001; Al-Mudimigh and Zairi, 2001; Schneider, 1999; Al-Mudimigh et al, 2001; Ngai et al, 2008; Nah et al, 2001; Alshamlam and Al-Mudimigh, 2011; Jing and Qiu, 2007). Due to change management system the old/ legacy systems is replaced with the new technologies and have more effective than the traditional systems. Although ERP system implementation is very expensive and high budget consuming but the success of its implementation enhance the efficiency of organizations. (Schneider, 1999) described that ERP projects are often expensive and almost half of the ERP project failed to achieve the target of promised benefits. It is evident that ERP systems implementation is based on the new technologies and complex in nature, so it is an intricate job for the potential users to handle and operate the new system effectively. Therefore, the user objection towards the new system in ERP implementation based on change management systems is common and observed almost everywhere.

ERP system is the collection of several modules with a single and integrated common database that help to integrate the business processes in the entire firm and also help to provide the main organizational behaviors in product and services (Aladwani 2001). Change management system is a system of tools, processes and principals that understanding the employees behavior and organizational transition from one state to another state throughout the ERP implementation for the success of an organization to achieve the goal (Al-Mudimigh and Zairi, 2001). ERP systems implementation change the whole setup of an organization by implementing a new and advance system therefore it need to be manage very carefully to achieve all the benefits of ERP systems (Al-Mudimigh et al, 2001). Therefore, it is required to introduce the change management system to the user to avoid the resistance of users towards the new system (Kemp and Low, 2008).

II. BACKGROUND STUDY

Change management for a successful ERP implementation is known to be a significant factor (Jing and Qiu, 2007; Zhang et al, 2003; Kuruppuarachchi et al, 2002; Masa'deh and Altamony, 2012; Delone & McLean, 1992; Summer, 1999; Kemp and Low, 2008). From the literature it is clear that change management provide the user to introduce with the new system and avoid the user resistance towards the new system and persuade the user behavior towards change (Masa'deh and Altamony, 2012; Kemp and Low, 2008). To adopt the new system and get all the expected benefits, enterprise necessarily needed to use the change management strategies and process (Ahmed et al, 2006; Kim et al., 2005). (Ahmed et al, 2006) further explained based on (Summer, 1999) that failure to ERP implementation observed in several companies due to no serious consideration on soft issues such as business processes and change management. Some researchers pointed out that in some organizations the employees and the management people don't want to implement the ERP system because they believe that the traditional manual system is comfortable to the management and the employees (Kuruppuarachchi et al, 2002) so that the change management is having no values to be implemented in the ERP system (Al-Nafjan and Al-Mudimigh,

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

2011). On the other hand ERP integrate different parts such as shared knowledge cost reduction and business process management improvement of an organization (Alshamlam and Al-Mudimigh, 2011).

ERP system implementation is risky, time consuming and expensive task (Alballa and Al-Mudimigh, 2011). In companies survey it is concluded by the (Jing and Qiu, 2007) that 44% companies reported that they had spent four times of the software license as much on the ERP implementation (Alballa and Al-Mudimigh, 2011; Jing and Qiu, 2007). The evidence of failure that is growing towards the ERP packages is due to fitting of organizational and national cultures that lead the projects which are expensive and late in delivery (Zhang et al, 2003). Furthermore, explained by (Zhang et al, 2003) "two measures are identified as indicators of the dependent variable for the success of ERP system implementation. The success of information system can be measure with six dimensions which are: system quality, information quality, use, user satisfaction, individual impact, and organizational impact" (Delone & McLean, 1992).

Resistance to change in ERP implementation is a challenging and never ending task. (Al-Nafjan and Al-Mudimigh, 2011) described in the case study that during ERP system implementation different change management strategies were adopted to deal with the resistance in change by communicating the ideas with the users to understand the logic of change (Summer, 1999). To overcome the resistance problem, the top management has to provide sufficient resources and clear vision of the benefits of ERP to the middle managements at each department (Alballa and Al-Mudimigh, 2011).

III. CHANGE MANAGEMENT STRATEGIES FOR SUCCESSFUL ERP IMPLEMENTATION

The primary concern of change management is to deal with the people challenges (Aladwani, 2001) and it is evident from the previous work in the literature that soft issues (people challenges) are more difficult than the technological issues (Aladwani, 2001; Alballa and Al-Mudimigh, 2011). One of the main issues in change management is the user resistance to the new system (Jing and Qiu, 2007). (Alshamlam and Al-Mudimigh, 2011) pointed out that the user resistance towards a particular problem should be understand and investigated that how they resist the new system. Furthermore, they explained that some users might be worried about their jobs in the organization or some user may have lack of technical skills for the new system whether to use the new system effectively or not.

Project champion in project management is a critical factor and has a vital role in successfully managing the change because the champion has a strong influence on the change process in the entire organization (Van Hau and Kuzic, 2010). (Ngai et al, 2008) described that successful and effective implementation of ERP system needed change management strategies and well build-up culture. Furthermore, to balance the change and user resistance, user training and education is necessary to educate them for the new system through user training systems so that to understand the new application and business processes of ERP system during their work in the organization (Legare, 2002). (Aladwani, 2001) further explained that feedback effective communication is an important factor in change management to understand the exact problem of user resistance towards the ERP system. Effective communication between the top management to their workers is mandatory to create the awareness in ERP system and keep them up-to-date of ERP system's benefits of ERP system (Alshamlam and Al-Mudimigh, 2011). Although the statistical figure only 41.6 % reported their usage of feedback through effective communication has been observed in successful projects in their change management strategy (Van Hau and Kuzic, 2010).

Cost minimization is another strategy of change management system (Aladwani, 2001). (Aladwani, 2001) based on (Porter, 1985) described a low cost strategy that can be used for a competitive business environment for the organization. Further they extended that this strategy is a useful suggestion for ERP where the new system is adopted by the user with the management approval and the user adoption cost should be minimal (Aladwani, 2001; Porter, 1985).

Readiness for change is important strategy in ERP implementation because when the employees/users are agree to change so it is a good sign for the top management to implement the new system and the system will be more useful with the readiness of employees willing (Alshamlam and Al-Mudimigh, 2011; Kwahk and Lee, 2008; Deloitte, 2005; Nah et al, 2001).

Strategies used in change management system have been summarized from different sources in the literature as shown in table 1. The order of the strategies is in descending order based on the sources collected from the literature.

TABLE I. CHANGE MANAGEMENT STRATEGIES

No	Strategies	References	No. of Source
1.	for Resistance	(Porter, 1985), (Aladwani, 2001), (Deloitte, 2005), (Kuruppuarachchi et al, 2002), (Alballa and Al-Mudimigh, 2011), (Nah et al, 2001), (Summer, 1999), (Ahmed et al, 2006), (Jing and Qiu, 2007), (Alshamlam and Al-Mudimigh, 2011), (Kwahk and Lee, 2008).	11
2.	for Soft Issues	(Aladwani, 2001), (Alballa and Al-Mudimigh, 2011), (Summer, 1999), (Ahmed et al, 2006), (Jing and Qiu, 2007), (Alshamlam and Al-Mudimigh, 2011).	6
3.	Readiness for Change	(Alshamlam and Al-Mudimigh, 2011), (Kwahk and Lee, 2008), (Deloitte, 2005), (Nah et al, 2001).	4
4.	Effective Communication	(Aladwani, 2001), (Alshamlam and Al-Mudimigh, 2011), (Van Hau and Kuzic, 2010).	3

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

5.	for Change in Processes	(Van Hau and Kuzic, 2010), (Ngai et al, 2008), (Legare, 2002).	3
6.	Cost minimization	(Aladwani, 2001), (Porter, 1985).	2
7.	User Friendly Culture	(Ngai et al, 2008).	1

IV. Change Management Processes for Successful ERP Implementation

The section of the paper is more concern about the tools, processes and methods that are using in ERP system implementation. From the literature, the authors have mentioned different tools with different names that were used during the change management in ERP implementation (Alshamlam and Al-Mudimigh, 2011) some called it change management activities (Al-Mashari, 2002) while some mentioned with change management factors (Deloitte, 2005). (Al-Mudimigh et al, 2001) described that change management is a set of tools, processes and activities that change the organizational setup from the current state to the expected future state by implementation of ERP systems to achieve the desired goals of an organization.

During change in the system effective communication is an important transition in ERP implementation. According to (Maditinos et al, 2012) effective communication is a trustworthy relationship between the external consultant and the employees of an organization. Moreover extended that the strong communication between the bodies, the more understanding between them (Fleck, 1993), and the less communication of consultant to users may lead to destabilize the ERP system implementation (Wang and Chen, 2006; Maditinos et al, 2012).

To ensure effective change in the organizations, user training and other important tools should be adopted for successful system implementation (Al-Mashari, 2002). Training and re-skilling of the IT personnel is an important initiative of change management (Aladwani, 2001). Employees training aims not only to understand and introduce the changes made to the system or how the new system will be operated but will help the users to understand that how new system change business processes of an organization (Alshamlam and Al-Mudimigh, 2011; Al-Mashari, 2002; Aladwani, 2001). And for that reason a helpdesk support system should be implemented to aware the employees about the changes and post implementation of the system (Nah et al, 2001).

Processes of change management system have been summarized from different sources in the literature as shown in table 2. The order of the processes is in descending order based on the sources collected from the literature.

TABLE II. Change Management Processes

No	Processes	References	No. of Source
1.	User Training	(Al-Mashari, 2002), (Alshamlam and Al-Mudimigh, 2011), (Aladwani, 2001), (Nah et al, 2001), (Porter, 1985), (Deloitte, 2005), (Kuruppuarachchi et al, 2002), (Alballa and Al-Mudimigh, 2011), (Summer, 1999), (Ahmed et al, 2006), (Jing and Qiu, 2007), (Kwahk and Lee, 2008),	12
2.	Tools and Methods	(Al-Mashari, 2002), (Deloitte, 2005), (Alshamlam and Al-Mudimigh, 2011), (Al-Mudimigh et al, 2001)	4
3.	Helpdesk	Alshamlam and Al-Mudimigh, 2011), (Al-Mashari, 2002), (Aladwani, 2001), (Nah et al, 2001)	4
4.	Communication	(Maditinos et al, 2012), (Fleck, 1993), (Maditinos et al, 2012).	3

V. Proposed Model

A. Methodology

Methodology of this paper is to discuss the change management strategies, techniques, processes and the success and failure of ERP system implementation. Moreover, the methodology has been extended with proposed model to strengthen the contents of change management systems. Proposed model declares the impact of change management on ERP system implementation as shown in figure 3. The focus of this paper is more concern on the user reactions against the new system and described in the proposed model.

B. Model contents

The contents of the proposed model includes: Change management strategies, change management processes, change management tools, user reactions, user training, user friendly culture, the role of consultants, and top management.

C. Model Descriptions

As discussed in the previous sections of this paper that ERP systems are very complex and expensive with high budget and more importantly that is very risky during implementation. Due to complex nature of ERP systems based on change management some important critical factors come in front in the implementation process. User reaction is one of the critical factors of change management and can lead the implementation process to failure if not considered seriously. The reaction against change rise when the strategies, processes and/or techniques of the system change from the existing setup to new system.

In this paper we have proposed a change management system success model for the successful ERP implementation

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

as shown in figure 3. In the proposed model the user reactions against change in change management is an important and critical factor as shown in figure 1.

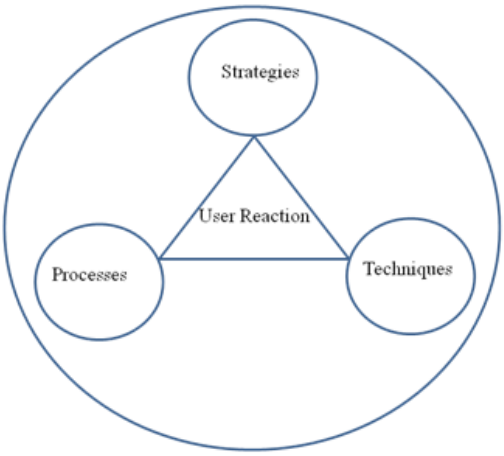


Figure 1. User reaction against change

This is the human nature that will never accept any change/complexity in the system against the systems they are using for years. The user reaction towards the new system is one of the three reasons:

- Lack of education: In some organizations the employees are hiring only for the purpose to input the data to the system and they don't care of the people with high education because the low educated people is hiring with minimum salaries.
- Lack of computer skills: The low educated people always having no expertise in computer programming and having no idea how to use the new applications because they learn some basic packages for their job survival and such people will never accept the complexity in the system to be implemented in their organization for the sake of their employment.
- Comfortable with the traditional system: From the literature, some authors mentioned that in many organizations the employees and even the managers don't like to disturb their old system because they are comfortable with the old/traditional system.

On the other hand, due to the rising benefits of ERP system implementation with advanced technologies is an important initiative for the organizational development. So the question arises that how to handle the user reaction? The proposed model aims to handle the user reaction against the changes made to the new system by providing them with effective training and technical education as suggested by most of the authors in the literature. Training and education help the user to introduce with the new system and understand them with the benefits of ERP system for the organizational outcomes. To deal with the user against their reactions, consultant needs to help them in understanding the new techniques and strategies of the implemented system by providing them training and education and user friendly culture to ensure the information flow between them is effectively operating for the success of organization, the scenario is shown in figure 2.

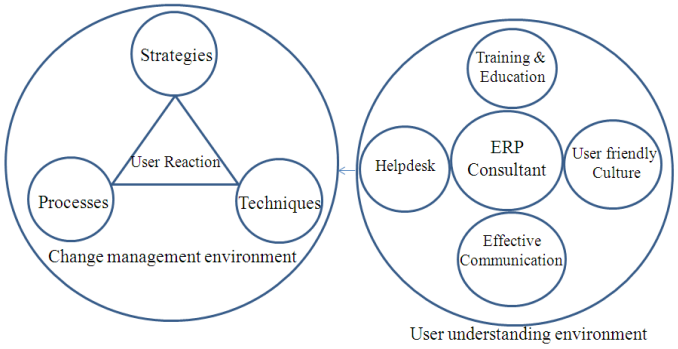


Figure 2. User training and education to understand the new system

After the system implementations, the consultants have a dominant role which is consisted of four important factors including Training and education, user friendly culture, establishment of helpdesk support center, and effective communication to handle the user reaction against the new system as shown in figure 2. Effective communication between the senior consultant and IT personal help the users to understand the complexity in new system and aware the personnel about the new changes for their convenience. Establishing a helpdesk support center is another important initiative to grip with the user reaction and ensure the user approval for further advancement in the system and solve their problem on time for their convenience. A user friendly culture is another important factor in ERP system implementation from the top management to provide them with cooperative and user friendly culture. It will help the individual to share their own knowledge for the organization development. The Top management is needed to provide the users/employees some encouragement and motivations such as incentives and rewards for their performance. The incentives and rewards system in the organization motivate the individuals involved in the system to improve their efficiency towards their potential work on the system for the betterment of organizational outcomes such as product and services and the company's future development. Once the reaction challenge handle and all other factors which are not the scope of this paper succeeded, the system will operate accordingly and the organization will get benefited as the promised befits of ERP outcome.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

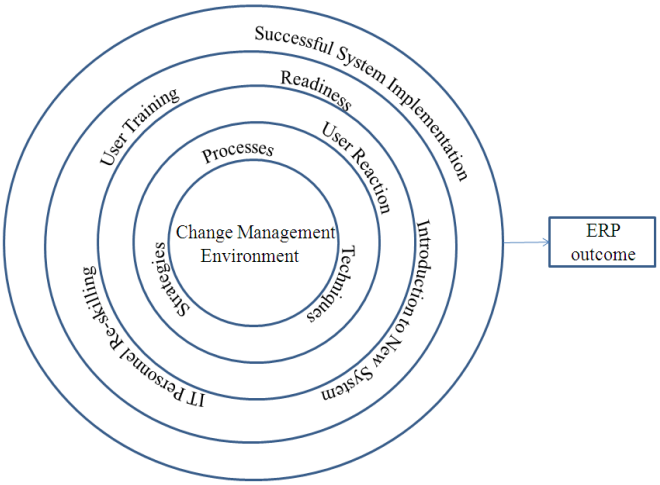


Figure 3. Proposed change management model for successful ERP implementation

VI. Critical Discussion

In this paper we discussed change management strategies, techniques, processes and the success and failure of ERP system implementation. A detailed and thorough study survey has been conducted and analyzed with all approaching strategies and promising techniques which are discussed in the above sections of this paper. The data has been collected from different sources including articles, case studies, books and online resources. The strategies discussed above clarify the vision of the new system implementation that is to be started with well planed strategies and to avoid all the reactions against the new system. The techniques and processes help the bodies involved, to understand the system and share their own expertise to achieve the company's goal with the successful ERP system implementation. User reaction towards change has been discussed above and it is known to be the critical issue in changing of the existing system. To cope with the reactions challenge towards the new system, user training and education is an important initiative in change management to provide them with technological skills and basic computer education by arranging training workshops for the users. User reaction towards change is a human natural obstacle that hurdles the way of changes in the existing system because they are using the existing system for years. To handle the issue of reactions, an incessant communication between the consultants and the IT personnel is necessarily required to aware them with the advancement in technology and the changes made to system. An effective communication support system (helpdesk) is an additional supplement that should be developed by the organization to update the users about the new processes and changes made to the system. The system will be more efficient if the user approve the changes made to the system and understand the benefits of ERP system. Once the system implemented successfully with the user approval and understanding, the users will share their own expertise for the best performance of organizational development such as product and services.

Different authors have different views regarding change management system but most of them agreed on some common concepts, processes and techniques. In the literature, almost all the authors described the issue of user resistance and they all agreed that user training and communication between consultant and end users is needed to cope with this challenge. Some authors pointed that user reaction towards change arise due to no sufficient IT or computer skills and they are worries about their jobs while other described that some people are technology enemy (Alshamlan and Al-Mudimigh, 2011), and they never accept the changes in the system and some other authors explained that some people don't like to disturb their system and they are feeling comfortable with their old systems (Kuruppuarachchi et al, 2002; Al-Nafjan and Al-Mudimigh, 2011). On the other hand some authors described that ERP systems are expensive and consuming with high budget so therefore it is out of their budget to implement the change in their setup. Although, The ERP system has installed in many organizations with successful outcomes exactly what was promised for the organization and they are running the system efficiently. The people related to the system sharing their own expertise and creating new knowledge for the system and implement it for the success of organizations.

As shown in figure 3, Change management systems is successful if the issue of user reactions is handle with strategies and provide them with training and educate them about the new technologies used in the new implemented system. Introduce the new system and aware them about the benefits of ERP systems for the company outcomes. Re-skilling of IT personnel is another strategy to train them for the new system. Moreover, the employees should be attracted with some motivational resources towards the new system by providing them with user friendly culture in the organization to share their own tacit knowledge for the improvement of company's performance. In a cooperative culture the top management is needed to encourage their employees with incentives and rewards so that the reaction towards the new implemented system is minimal.

VII. Conclusion

ERP is an integrated and business information software system that attracts the attention of business organizations in order to improve their business processes and achieve the company goals. For that reason, the ERP system should be more concern on the business processes change than the technical change during installation. ERP system provides a user friendly culture due to which the individual can share their own capabilities for the organization future development.

Change management in ERP system implementation is a challenging task because the barriers in the form of soft issues come out when the changes occur to the existing system. Change management believes on change, therefore the user reaction towards the new system is obviously not new and can be expected at any stage during ERP implementation process. To minimize the reactions against the new system user training and education is mandatory initiation to introduce the new system to the users so that they will understand it and use it

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

effectively. Effective communication between the consultant and IT personnel is another solution to cope with this challenge in change management.

The proposed model presented in this paper corroborates the impact of change management on ERP systems implementation. The aim and scope of the presented model is to handle all the possible barriers that hurdle the implementation process and provide solutions to users and ensure the success of system implementation.

From the discussion based on literature, it is concluded that ERP system implementation for the business organizations and business processes management is an important initiation. Although the system is complex and high budget consuming but the outcomes of the system is benefited for the organization's product and services and competitiveness in the market.

REFERENCES

[1] Aladwani, A. (2001), "Change management strategies for successful ERP implementation" Business Process Management Journal, Emerald Group Publishing Limited 7 (3): 266-275(210).

[2] Schneider, P. (1999), "Wanted: ER People Skills", CIO Magazine, 12 (10), pp. 30-37.

[3] Al-Mudimigh. A., Zairi. M., Al-Mashari. M., (2001), "ERP software implementation: an integrative framework", European Journal of Information Systems. Vol.10, pp 216–226.

[4] Kemp. M. J., Low. G. C., (2008), "ERP innovation implementation model incorporating change management", Business Process Management Journal 14(2): 228-242.

[5] Ngai, E. W. T., Law. C. C. H., Wat, F.K.T. (2008), "Examining the critical success factors in the adoption of enterprise resource planning", Compute Industry (Ind) Volume 59, Issue 6, pp 548–564.

[6] Almudimigh, A., Zairi, M. (2001), "ERP systems implementation: A best practice perspective and a proposed model." The European Centre for Total Quality Management (ECTQM), Report No. R-01-01.

[7] Van Hau. T. T, Kuzic. J., (2010), "Change Management Strategies for the Successful Implementation of Enterprise Resource Planning Systems", Second International Conference on Knowledge and Systems Engineering, pp 178-182

[8] Legare. T. L., (2002), "The Role of Organizational Factors in Realizing ERP Benefits", Information Systems Management, Volume 19, Issue 4, pp 21-42.

[9] Alballaa. H, Al-Mudimigh. A. S., (2011), "Change Management Strategies for Effective Enterprise Resource Planning Systems: A Case Study of a Saudi Company", International Journal of Computer Applications, Volume 17– No.2, pp 14-19.

[10] Jing, R., Qiu, X. (2007), "A Study on Critical Success Factors in ERP Systems Implementation." Service Systems and Service Management, 2007 International Conference 1-6.

[11] Porter, M. (1985), "Competitive Advantage: Creating and Sustaining Superior Performance", the Free Press, New York, NY.

[12] Kwahk. K., Lee. J., (2008), "The role of readiness for change in ERP implementation: Theoretical bases and empirical validation", Information & Management Journal, Vol.45, pp 474–481.

[13] Deloitte., (2005), "ERP Change Management Survey", The Gullup Leadership Institute.

[14] Nah. F. F., Lau. J. L., Kuang. J., (2001), "Critical factors for successful implementation of enterprise systems", Business Process Management Journal, Vol. 7, pp. 285-296

[15] Masa'deh. R., Altamony. H., (2012), "A Theoretical Perspective on the Relationship between Change Management Strategy and Successful ERP Implementations", Research Journal of International Studies, Euro Journals Publishing, Inc. 2012, pp 141-154.

[16] Ahmed. Z. U., Zbib. I., Arokiasamy. S., Ramayah. T., Chiun. L. M., (2006), "RESISTANCE TO CHANGE AND ERP IMPLEMENTATION SUCCESS: THE MODERATING ROLE OF CHANGE MANAGEMENT INITIATIVES", Asian Academy of Management Journal, Vol. 11, No. 2, pp 1–17.

[17] Summer. M., (1999), "Critical success factors in enterprise wide information management systems projects", Proceedings of the Americans Conference on Information Systems (AMICS), Milwaukee, WI, 232–234

[18] Kim. Y., Lee. Z., Gosain. S., (2005), "Impediments to successful ERP implementation process", Business Process Management Journal, 11(2), 158–170

[19] AL-NAFJAN, A. N., Al-MUDIMIGH. A. S., "THE IMPACT OF CHANGE MANAGEMENT IN ERP SYSTEM: A CASE STUDY OF MADAR", Journal of Theoretical and Applied Information Technology, Vol. 23 No 2, pp 91-97.

[20] Kuruppuarachchi, P., Mandal, P., Smith, R., (2002), "IT project implementation strategies for effective changes: A critical review", Logistics Information Management (2), 126-137.

[21] Zhang. L., Lee, M. K. O., Zhang. Z., Banerjee. P., (2003), "Critical Success Factors of Enterprise Resource Planning Systems Implementation Success in China", Proceedings of the 36th Hawaii International Conference on System Sciences.

[22] Al-Mashari. M. A., (2002), "Implementation ERP through SAP R/3: A Process Change Management (PCM) perspective", J. King Saud Univ., Vol.14, Comp. & Info. Sci, pp 25-38.

[23] Maditinos. D., Chatzoudes. D., Tsairidis. C., (2012), "Factors affecting ERP system implementation effectiveness", Journal of Enterprise Information Management Vol. 25 No. 1, pp. 60-78.

[24] Fleck. J., (1993), "Configurations: crystallizing contingency", International Journal of Human Factors in Manufacturing, Vol. 3 No. 1, pp. 15-36.

[25] Wang, E., Chen. J., (2006), "Effects of internal support and consultant quality on the consulting process and ERP system quality", Decision Support Systems, Vol. 42, pp 1029-41.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

SECURING AODV WITH AUTHENTICATION MECHANISM USING CRYPTOGRAPHIC PAIR OF KEYS

K.Suresh Babu
Research Scholar
School of IT
JNT University Hyderabad, India.

K.Chandra Sekharaiah
Professor in CSE
School of IT
JNT University Hyderabad, India

***Abstract* -- Mobile Ad Hoc Networks (MANETs) is characterized by self–organizing capability, dynamically configurable infrastructure and multihops. Of late, MANETs form emerging state-of-the-art networking technology faster. The routing protocol plays an important role in it overall operation of MANETs. AODV is one of MANET routing protocol. In this paper, the vulnerabilities in MANETs and security flaws in AODV are discussed. A new security mechanism for securing AODV with message digest authentication using a pair of keys (public key cryptography) is proposed and implemented in NS - 2 simulator.**

Keywords – Self-organizing; multihops; authentication; public key

I. Introduction

Mobile Ad Hoc Networks (MANETs) is characterized by self–organizing capability, dynamically configurable infrastructure and multihops. Of late, MANETs form emerging state-of-the-art networking technology faster[1]. Mobile Ad Hoc Network (MANET) technology is designed for the establishment of a network anywhere and anytime, without any fixed infrastructure. The mobility and dynamism are the important features of this networks. To support these features the required mechanism is embedded in the mobile nodes itself. As, the MANETs are self-configured networks and allow ubiquitous service access, anywhere, anytime without any fixed infrastructure they can have several types of applications like rescue operations, military, law enforcement and security operation, home network and conferencing .

Mobile ad hoc networks (on-the- fly) are characterized by lack of infrastructure .Nodes in a network are free to move and organize themselves in a arbiter fashion. Communication between two nodes may have multiple links and heterogeneous radio, and can operate in a stand-alone fashion, well suited in a situation where infrastructure is unavailable or cost effective, time effective and also be used in crises management service applications. MANET has received good attention because of its self-configuration and self-maintenance capability

Security is always the top issue in computing. It maintains the order of a system or a network of systems. Any unauthorized action (i.e., altering the system files) might cause failures or loss of valuable data. So, security must be taken seriously during the design and analysis of secure systems.[3]. Now a day's patterns of security are developed in the software development. Security patterns can and should be applied to develop secure systems. These security patterns help the systems to be more secured. The patterns should be applied at each and every stage of software development life cycle.[2]. Security is challenging task in MANET because of its characteristics.

The security of the MANETs has to be discussed at routing protocols only. Beacuase routing protocols are the very important in the overall operation of MANETs. The existing MANET routing protocols are emphasizing more on security. In this paper, we propose some modifications existing AODV MANET routing protocol so that we can embed security components to it.

The main aim of our routing protocol is to distribution of public key to other nodes, and to provide authentication integrity and non repudiation in MANET using AODV message with extension of fields in already existing AODV protocol. The rest of the paer is organized as follows, section 2 details the security attacks on MANETs, section 3 deals with AODV routing protocol features, section 4 enlightens the security flaws in AODV. In section 5 the proposed work is presented, section 6 gives the simulation results and section 7 gives the conclusion.

II. Security Attacks on MANETs

The characteristics like dynamically varying network topology [2], imprecise state information, lack of central coordination, hidden node problem, limited resource and insecure medium. Each node in a MANET act as a host and router means it forms a peer to peer network. It is a fundamental vulnerability and there is no clear line of defense in security design and no well define place to deploy security solution. Heterogeneous nodes are present in MANET, which leads to device physical capture [5] attack. The computational capacity of nodes is constrained, they hardly perform intensive task like cryptographic computation.

In MANET communication is due to single hop through link layer protocol and multi hop through network layer protocol .These protocols typically assumes that all node in a network are cooperative in coordination process. But this assumption is unfortunately not true in hostile environment .Cooperation is assumed but not enforced in MANET ,malicious attacks can easily

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

disrupt network operation by violating protocol specifications. The network layer operation in MANETs are routing and data packet forwarding. But both are vulnerable to malicious attacks, leading to various types of malfunction in network layer[11]. The various types of attacks are:

i. Eavesdropping: Eavesdropping is interaction and reading of message by unintended receiver .In MANET nodes share wireless medium, majority of wireless communication use the RF spectrum and broadcasting .So signals broadcast over airwaves can be easily interacted with receivers tuned to the proper frequency. Thus, message transmitted can be eavesdropped, and fake messages can be injected into network.

ii. Route Discovery Attacks: In MANET malicious routing attacks that target the 1) Routing Discovery 2) Route Maintenance phase by not following the specification of routing protocols.

iii. Routing Table Overflow Attack: In proactive routing protocols, updation of routing information is done periodically and it discovers the routing information before it actually needed. The attacker tries to create enough routes to prevent new route from being created. Attacker simply creates excessive route advertisements to overflow the victims routing table [11].

iv. Routing cache Poisoning Attack: Route cache poisoning in DSR [12]. This is a passive attack that can occur in DSR due to promiscuous mode of updating routing table which is employed by DSR. This occurs when information stored in routing table at routers is deleted, altered or injected with false information.

v. Attacks at routing maintenance Phase: In MANET, there are attacks that target the route maintenance phase by broadcasting false control messages, which cause invocation of the costly route maintenance or repairing operation.

vi. Attacks at Data forwarding Phase: Some attacks target data packet forwarding functionality at the network layer. Here, malicious nodes cooperate with routing protocol in route discovery phase, but not in the forwarding phase, they do not forward the packet consistently according to routing table.

vii. End to End Attack: Similar to TCP protocol in the internet, the mobile node is vulnerable to classic "SYN" flooding attack or Session Hijacking attack. In TCP session hijacking attack, the attacker spoof the victim IP address, and determine the correct sequence number that express by the target and then perform DOS attack on the victim

III. Ad hoc On-Demand Distance Vector (AODV) Routing Protocol

Ad hoc On-Demand Distance Vector (AODV) is a well-known routing protocol which is based on distance vector for MANETs. It is on demand *reactive* protocol, in which nodes in the network exchange routing information only when a communication needs to take place, and keep this information up-to-date only as long as the communication period. It won't give the total node view each node know about its neighbor node[5]. When a node wants to communicate with other node, it starts a *route discovery* process in order to establish a route towards the destination node by using route request packet. Therefore, it broadcast a route request message (RREQ).

<source add, source sequence _#, broadcast id, destination add, destination sequence_#, hop count>

The pair of <source address, broadcast id> uniquely identify RREQ .broadcast id incremented whenever the source issues a new RREQ. AODV uses *sequence numbers* in order to identify fresher routing information. Each node maintains its own sequence number, incrementing it before sending a new RREQ or RREP message. Whenever Neighboring nodes receive the RREQ, they increment the hop count, and broadcast the message to their neighbors[6]. The goal of the RREQ message is to find the destination node, but it also has the side effect to make other nodes learn a route towards the source node (the "Reverse route"): a node that has received a RREQ message with source address (let S) from its neighbor (let N) knows that it can reach S through N, and records this information in its routing table along with the hop count (i.e., its distance from node S following that route). Whenever RREQ message reach the destination node, then destination react with a route reply message (RREP).

<Source address, destination address, destination sequence_#, hop count, life time >

The RREP is sent as a uncast, using the path towards the source node established by the RREQ. Similarly to what happens with RREQs, the RREP message allows intermediate nodes to learn a route towards the destination node (i.e., the originator of the RREP). Therefore, at the end of the route discovery process, packets can be delivered from the source to the destination node and vice versa.
AODV[7] uses RERR to notify errors to nodes. AODV makes use of HELLO messages periodically to find link failures to nodes that it considers as its immediate neighbors. When a link failure is detected for a next hop of an active route a RERR (Route Error) message is sent to its active neighbors that were using that particular route. The key **vulnerabilities** [6] present in the basic AODV routing protocol are: 1) Deceptive incrementing of Sequence Numbers 2) Deceptive decrementing of Hop Count

IV. Security flaws of AODV

AODV does not provide any security mechanism so any malicious node can perform any attack. The protocol does not talk anything about the security services like authentication, confidentiality. A malicious node can perform the following: 1) Malicious node can send RREQ or RREP packet as it is an authorized node either source or destination. 2) Malicious node selectively not forwarding RREQ and RREP. 3) Malicious node intentionally increases or decreases the hop count in RREQ packet.4) Malicious node intentionally can change the time out field in RREP packet.

V. Proposed Work

In our work we have made some modifications to the existing AODV algorithm. The packets of the AODV have been changed. We have implemented the encryption and authentication mechanisms in this work. The RSA algorithm is used for encryption and MD5 authentication is used for achieving authentication.

It is assumed that every node has to maintain key pair i,e. private key and public key and also every node has to maintain two counters i.e., sequence number and broadcast _id. Proposed system uses a message digest with public key to secure AODV communication .It calculates message digest using MD5 to all the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

fields of an AODV message in addition with public key. Calculated Message digest value transmitted along with AODV message.

The RREQ packet of the proposed system, has one field extension in RREQ message i.e., public key and message digest

<source add, source sequence _#, broadcast id, destination add, destination sequence_#,
Hop count, public key, encrypted message digest>

TYPE	J	R	G	RESERVED	HOP COUNT
RREQ ID					
DESTINATION IP ADDRESS					
DESTINATION SEQUENCE NUMBER					
SENDER IP ADDRESS					
SENDER SEQUENCE NUMBER					
PUBLIC KEY OF SENDER					
ENCRYPTED MESSAGE DIGEST					

Fig. 1: Proposed RREQ Packet

The proposed RREP packet contain one field extension i.e., encrypted message digest

<Source address, destination address, destination sequence_#, hop count ,life time ,public key ,encrypted message digest>

TYPE	R	A	RESERVED	PREFIX SIZE	HOP COUNT
DESTINATION IP ADDRESS					
DESTINATION SEQUENCE NUMBER					
ORIGINATOR IP ADDRESS					
LIFE TIME					
ORIGINATOR PUBLIC KEY					
ENCRYPTED MESSAGE DIGEST					

Fig. 2: Proposed RREP/RERR Packet

Secure AODV Algorithm

1) Node which generates RREQ or (RREP / RERR), it has to calculate the message digest using MD5 all the fields of proposed RREQ message excluding message digest field.
2) Encrypt the message digest using private key
3) It has to put encrypted message digest value in the message digest field in RREQ message, and broadcast it.
4) The node receives the RREQ or (RREP / RERR) packet.
5) If it is destination then it collects the public key from the RREQ message fields, and decrypt the encrypted message using the public key let (say X). It calculates the message digest excluding encrypted message digest field with making hop count as zero say (MD1) of received RREQ packet let (say Y) , do the verification X= =Y.
6) Else, i.e., it is not a destination means inter mediatory node. Then rebroadcasting RREQ or forwarding RREP/RERR node with increment of hop count in the field of message.

VI. Simulation and Security analysis

We have used Network Simulator Version-2 (NS2) [8, 9] to simulate our proposed protocol. We have successfully implemented message digest with RSA encryption mechanism to secure AODV routing protocol using NS- 2.35 [8, 9] on Linux operating system requirements without consuming much power of nodes. This mechanism gives better performance compared with existing AODV protocol in throughput, Delay. The main aim of simulation is to prove proposed mechanism is properly securing AODV with all security aspects. For simulation, we have considered 3 different aspects i.e., number of nodes with varying the misbehavior nodes count.

No. of Nodes	10, 20, 30
Area Size	600 * 600
Mac	802.11
Radio Range	250m
Simulation Time	100 sec
Traffic Source	CBR
Packet Size	512
Mobility Model	Random Way Point
Speed	2,4,6,8,10 and 12 m/sec.

Fig. 3: Simulation Scenario

All network components of mobile nodes are considered their default values. (E.g. Link Layer, Interface Queue, Mac Layer etc.) Agent, Router and Movement traces are kept ON and Mac trace is kept OFF for all three mobile nodes. In below diagram shows our network scenario.

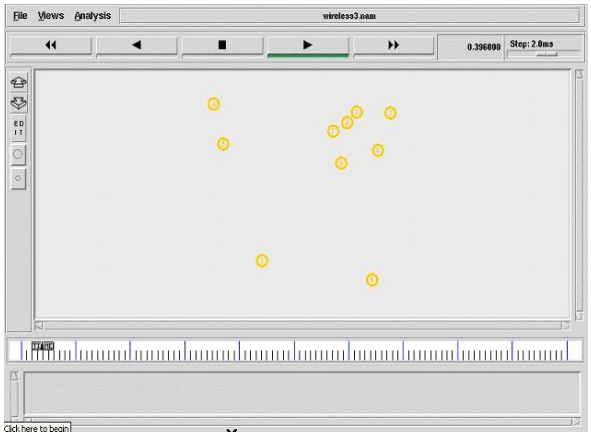


Fig. 4: Simulation of nodes in NS 2

Performance Metrics: We used the QoS parameters throughput and delay to evaluate the performance of our proposed protocol.

- Throughput: This is the ratio of packets received by the receiver to packets delivered by the sender (CBR packets delivered)
- Average end-to-end delay: This is the average of the delays incurred by all the packets that are successfully transmitted

It is been observed from the X graph that both throughput and delay have been improved with secured AODV.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Fig. 5: Graph showing how the proposed system protocol has shown a good decrease in delay when compared to the general AODV when there are misbehavior nodes.

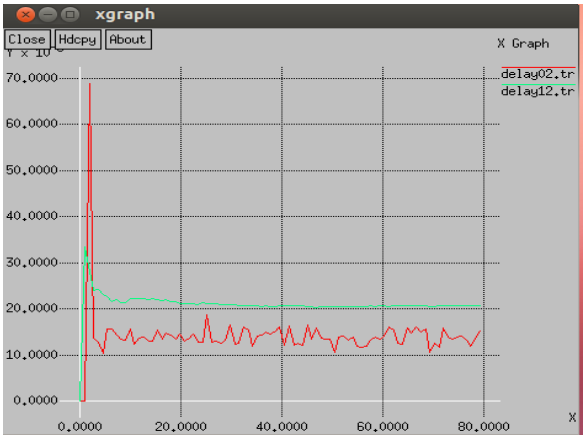


Fig. 6: Graph showing how the proposed system protocol has shown a good increase in throughput when compared to the general AODV when there are misbehavior nodes.

Security Analysis: Security analysis is also done and we could find that the following security services were achieved.

1) Authorization All the nodes have unique key pair to perform the hash function and encryption

2) Authentication All the nodes authenticated by their public key, if any malicious node want to authenticate in network then it has to gain public key pair of authorized node.

3) Non Repudiation Public key send along with the route computation avoid non repudiation.

4) Integrity Message is protected by tampering due to encrypt with the destination public key.

VII. Conclusion

In this paper we have discussed security issues in MANETs. A study of AODV routing algorithm is taken place and security flaws in the algorithm are learnt. To provide the security for the AODV protocol, we have proposed and implemented a secure mechanism to AODV protocol by using RSA encryption algorithm and MD5 authentication algorithms. Security services such as authorization, confidentiality, authentication, non-repudiation and integrity have been achieved. Finally, we conclude that the mechanism is secure and efficient. In our future work we want to implement the same mechanism on other MANET routing protocols.

VIII. References

[1] K.Suresh Babu, K.ChandraSekharaiah, *"Mobile Ad-Hoc Networks : A Novel Survey", International Conference On Advanced Computing And Communication Technologies For High Performance Applications, FISAT, COCHIN, September 24-26' 2008, Vol. 1, Page.262-269.*

[2] K.Suresh Babu, K.ChandraSekharaiah, **"***Security Patterns: State-of-Art Scenario" International Journal of Computer Science and Network Security(IJCSNS), ISSN: 1738-7906, April 2011, Vol. 11, No.4, Page.131-135.*

[3] K.Suresh Babu, K.ChandraSekharaiah, "System Security: A Survey", National Conference Proc. RESPOGRAF, ASTRA, 2008.

[4] K.Suresh Babu, K.ChandraSekharaiah, "*Issues Related to Routing and Security in Mobile Ad-Hoc Networks", January 2009, CI-4.7, International Conference on Systemics, Cybernetics and Informatics ICSCI-2009, January 07-10 2009.*

[5] Junaid Arshad, Mohammad Ajmal Azad, "Performance Evaluation of Secure on-Demand Routing Protocols for Mobile Ad-hoc Networks", *(2006) IEEE*, pp. 971-975.

[6] Asad Amir Pirzada, Chris McDonald, "Secure Routing with the AODV Protocol", *(2005) Asia Pacific Conference on Communication, Perth, IEEE*, p.p. 57-61.

[7] Perkins, Belding-Royer and Das, "Ad hoc on-demand distance vector (aodv) routing", *IETF RFC 3591, 2003.*

[8] Ns homepage - http://www.isi.edu/nsnam/ns/

[9] Ns manual - http://www.isi.edu/nsnam/ns/

[10] Ankita Gupta, Sanjay Prakash Ranga, "VARIOUS ROUTING ATTACKS IN MOBILE AD-HOC NETWORKS", Vol 2 Issue 4 July 2012, IJCCR, ISSN 2249-054X.

An Overview of Wireless Local Area Networks (WLAN)

Ibrahim Al Shourbaji

Computer Networks Department
Jazan University
Jazan 82822-6649, Saudi Arabia

Abstract

Wireless Communication is an application of science and technology that has come to be vital for modern existence. From the early radio and telephone to current devices such as mobile phones and laptops, accessing the global network has become the most essential and indispensable part of our lifestyle. Wireless communication is an ever developing field, and the future holds many possibilities in this area. One expectation for the future in this field is that, the devices can be developed to support communication with higher data rates and more security. Research in this area suggests that a dominant means of supporting such communication capabilities will be through the use of Wireless LANs. As the deployment of Wireless LAN increases well around the globe, it is increasingly important for us to understand different technologies and to select the most appropriate one .

This paper provides a detailed study of the available wireless LAN technologies and the concerned issues ,will give a brief description of what wireless LANs are ,the need of Wireless LAN ,History of wireless LAN , advantages of Wireless Networks ,with summarizing the related work on WLAN in academic area , Wireless LAN technologies , some risks attacks against wireless technologies , suggesting some recommendations to protect wireless LAN network from attack , Finally we propose some research issues should be focused on in the future. .

Keywords: *Wireless Networking, Security, 802.11 Standard, Network security,*

I. INTRODUCTION

Computer technology has rapidly growth over the past decade, Much of this can be attributed to the internet as many computers now have a need to be networked together to establish an online connection. As the technology continues to move from wired to wireless, the wireless LAN (local area network) has become one of the most popular networking environments.

Companies and individuals have interconnected computers with local area networks (LANs).The LAN user has at their disposal much more information, data and applications than they could otherwise store by themselves. In the past all local area networks were wired together and in a fixed location. Wireless technology has helped to simplify networking by enabling multiple computer users simultaneously share

resources in a home or business without additional or intrusive wiring.

The increased demands for mobility and flexibility in our daily life are demands that lead the development

2. What is a WLAN ?

To know WLAN we need first to know the definition of LAN, which is simply a way of connecting computers together within a single organization, and usually in a single site (Franklin, 2010).

According to Cisco report in 2000 wireless local-area network (WLAN) does exactly what the name implies: it provides all the features and benefits of traditional LAN technologies such as Ethernet and Token Ring without the limitations of wires or cables. Obviously, from the definition the WLAN is the same as LAN but without wires.

(Clark et al, 1978) defined WLAN as a data communication network, typically a packet communication network, limited in geographic scope.' A local area network generally provides high-bandwidth communication over inexpensive transmission media.
While (Flickenger, 2005) see it as a group of wireless access points and associated infrastructure within a limited geographic area, such as an office building or building campus, that is capable of radio communications. Wireless LANs are usually implemented as extensions of existing wired LANs to provide enhanced user mobility.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Wireless Local Area Network (WLAN) links two or more devices using a wireless communication method. It usually provides a connection through an Access Point (AP) to the wider internet (Putman, 2005).

This gives users the ability to move around within a local coverage area while still be connected to the network. Just as the mobile phone frees people to make a phone call from anywhere in their home, a WLAN permits people to use their computers anywhere in the network area.

In WLAN Connectivity no longer implies attachment. Local areas are measured not in feet or meters, but miles or kilometers. An infrastructure need not be buried in the ground or hidden behind the walls, so we can move and change it at the speed of the organization.

3. Why would anyone want a wireless LAN?

There are many reasons: (perm, 2000)

1- An increasing number of LAN users are becoming mobile. These movable users require that they are connected to the network regardless of where they are because they want simultaneous access to the network. This makes the use of cables, or wired LANs, impractical if not impossible.

2- Wireless LANs are very easy to install. There is no requirement for wiring every workstation and every room. This ease of installation makes wireless LANs inherently flexible. If a workstation must be moved, it can be done easily and without additional wiring, cable drops or reconfiguration of the network.

3- Another advantage is its portability. If a company moves to a new location, the wireless system is much easier to move than ripping up all of the cables that a wired system would have snaked throughout the building.
Most of these advantages also translate into monetary savings.

4. History of WLAN

(Negus & Petrick, 2009)
The wireless local area network (WLAN) is today everywhere device often taken for granted as a default interface for networked devices by users and manufacturers alike. But not very long ago, it was most definitely not so.

In the early 1990's WLANs found almost no success in selling to enterprise or campus environments as wired LAN replacements or enablers of mobility. The WLAN products of that day were far too slow, too expensive, too bulky, and too power hungry. Furthermore, mobile network connectivity was simply not yet a killer application. The "survivor" companies of that age were the ones who focused on adapting WLAN technology to specialty niches such as retailing, hospitality, and logistics.

Organizations that went after the "big" market of enterprise networking, and there were many that did, either went bankrupt or became largely scaled back divisions of large companies.

By the middle of the 1990's the WLAN industry had mainly consolidated into 4 players, But in the late 1990's the first significant market opportunity for WLANs emerged and it was quite unlike what the WLAN industry to date had largely envisioned.

The opportunity was the sharing of a broadband Internet connection within the home amongst multiple networked devices such as PCs initially, but inevitably also voice over Internet protocol (VoIP) phones, gaming consoles, media streamers and home automation appliances. Consumers, not enterprise IT managers, became the ones to choose what WLAN technology and products would achieve the de facto standard for the decade to follow.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Advantages of Wireless Networks

Wireless LANs designed to operate in license-free bands making their operation and maintenance costs less than contemporary cellular and PCS networks. The use of license-free spectrum, however, increases the risk of network security and in-band interference. The key advantages of wireless networks as opposed to wired networks are mobility, flexibility, ease of installation and maintenance, and reduced cost. (Aziz, 2003)

According to (Symantec , 2002) wireless LANs are less expensive and less intrusive to implement and maintain, as user needs change.
Simple implementation and maintenance, extended reach,, increased worker mobility and reduce total cost of ownership and operation.

Emerging Developments

Fundamental step forward in information theory, which first emerged during the time of the early development of wireless LANs, have now reached a level of maturity and acceptance that is allowing them to drive the quest for higher spectral efficiencies and data rates.

Another important development in wireless LAN technology is the emergence of mesh networking. Mesh networks have the potential to dramatically increase the area served by a wireless network. Mesh networks even have the potential, with sufficiently intelligent routing algorithms to boost overall spectral efficiencies attained by selecting multiple hops over high capacity links rather than single hops over low capacity links (Holt, 2005).

5-Wireless LAN Technologies

When making a decision about the best protocol or standard to use. We need to consider its features and our needs. Weight the features and compare the advantages and disadvantages of each one to make the final decision.

There are several wireless LAN solutions available today, with varying levels of standardization and interoperability.
Many solutions that currently lead the industry, IrDa, Bluetooth, HomeRF and IEEE 802.11. These technologies enjoy wider industry support and targeted to solve Enterprise, Home and public wireless LAN needs.

- **Infrared** (IrDa)

The appearance of portable information terminals in work and living environments is increase the introduction of wireless digital links and local area networks(LAN's).

Wireless LANs can use either radio frequencies or infrared light to transmit signals. While it is considerably cheaper to install infrared networks, as many devices already have infrared (IrDA) ports (Franklin, 2010).

Portable terminals should have access to all of the services that are available on high-speed wired networks. Unlike their wired counterparts, portable devices are subject to severe limitations on power consumption, size and weight. The desire for inexpensive, high-speed links satisfying these requirements has motivated recent interest in infrared wireless communication (Gfeller & Bapst, 1979).

Wireless infrared communications refers to the use of free-space propagation of light waves in the near infrared band as a transmission medium for communication (Carruthers, 2002).

The Infrared Data Association (IrDA) is another trade association, which defined standards for infrared communication for many years. It has some advantages; notably that it is cheap and there are many devices which already include infrared including most laptops and PDAs as well as some printers. Before the advent of radio frequency LANs people were building infrared LANs, with some success. (irda.org, 2011)

The wavelength band between about 780 and 950 nm is presently the best choice for most applications of infrared wireless links, due to the availability of low-cost LED's and laser diodes (LD's), and because it coincides with the peak responsively of inexpensive, low-capacitance silicon photodiodes (Rancourt,, 1993).

It provide a useful complement to radio-based systems, particularly for systems requiring low cost, light weight, moderate data rates, and only requiring short ranges (Carruthers, 2002).

However, this radiation cause problem relates to eye safety; it can pass through the human cornea and focused by the lens onto the retina, where it can potentially induce thermal damage (Kahn & Barry, 1997).

To achieve eye safety with an LD user can employ a thin plate of translucent plastic. such diffusers can achieve efficiencies of about 70%, offering the designer little freedom to tailor the source radiation pattern. Computer generated holograms (Smyth et al, 1995).

The primary goals in extending IrDA-Data's connection model were: (Williams, 1999)

- To enable devices to view each other to establish communication relationships uninhibited by the connection state of nearby devices.
- To enable an AIR device to establish communications with at most one IrDA 1.x device.
- For AIR devices to respect established connections with which they could interfere. This is a co-

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

existence requirement intended to ensure that AIR devices do not disrupt active connections

- **Bluetooth**

Bluetooth is an industry specification for short-range connectivity for portable personal devices with its functional specification released out in 1999 by Bluetooth Special Interest Group.
Bluetooth communicates on a frequency of 2.45 gigahertz, which has been set aside by international agreement for the use of industrial, scientific and medical devices (ISM) (Chandramouli, 2005). It is a worldwide license free band that any system can use (Goldsmith, 2004).

Using this band allows the Bluetooth protocol to become a standard around the world for interfacing devices together wirelessly.
Communications protocol developed to allow the devices using Bluetooth to transfer data reliably over their wireless network.

Bluetooth has a range of less than 10 meters. The range is increased when a scatternet is used because each unit only has to be within 10 meters of one other unit. The range can also be increased if the data is transmitted in a high power mode which offers transmissions up to 100 meters. Bluetooth also offers a cipher algorithm for security. This is most useful in the high power mode because when data is being transmitted further there is a greater possibility of an unwanted device receiving the network's data (Goldsmith, 2004).

- **HomeRF**

In early 1997, several companies formed the Home RF working group to begin the development of a standard designed specifically for wireless voice and data networking in the home. (Goldsmith, 2004). HomeRF is an open industry specification developed by Home Radio Frequency Working Group (Wireless Networking Choices for the Broadband Internet Home., 2001) that defines how electronic devices such as PCs, cordless phones and other peripherals share and communicate voice, data and streaming media in and around the home.

The development of this working group was motivated by the widespread use of the internet and the development of affordable PCs that can be used in most homes. This protocol allows PCs in the home to have greater mobility, providing a connection to the Internet, printers, and other devices anywhere in the home. With all this potential, many members of industry worked to develop the Shared Wireless Access Protocol-Cordless Access (SWAP-CA) specification (Goldsmith, 2004).

Unlike Wi-Fi, HomeRF already has quality-of-service support for streaming media and is the only

wireless LAN to integrate voice. HomeRF may become the worldwide standard for cordless phones. In the year 2001, the Working group unveiled HomeRF 2.0 that supports 10 Mbps (HomeRF 2.0) or more. (Chandramouli, 2005)

A network topology of the Home RF protocol consists of four types of nodes: Control Point, Voice Terminals, Data Nodes, and Voice and Data Nodes. The control point is the gateway to the public switched telephone network (PSTN) and the Internet. It is also responsible for power management of the network. A voice terminal communicates with the control point via voice only. A data node communicates with the control point and other data nodes. Finally, a voice and data node is a combination of the previous two nodes (Lansford, 2000).

- **IEEE 802.11**

The vendors joined together in 1991, first proposing, and then building, a standard based on contributed technologies. In June 1997, the IEEE released the 802.11 standard for wireless local-area networking (Cisco Wireless Lan standard report, 2000).

This initial standard specifies a 2.4 GHz operating frequency with data rates of 1 and 2 Mbps. With this standard, one could choose to use either frequency hopping or direct sequence. Because of relatively low data rates as, products based on the initial standard did not flourish as many had hoped (Chandramouli, 2005).

In late 1999, the IEEE published two supplements to the initial 802.11 standard: 802.11a and 802.11b (Wi-Fi). The 802.11a (Highly Scalable Wireless LAN Standard , 2002), standard (High Speed Physical Layer in the 5 GHz Band) specifies operation in the 5 GHz band with data rates up to 54 Mb/s (O'Hara, B. and Petrick, 1999).

The 802.11 WLAN standard allows for transmission over different media. Compliant media include infrared light and two types of radio transmission within the unlicensed 2.4-GHz frequency band:
frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS).
Spread spectrum is a modulation technique developed in the 1940s that spreads a transmission signal over a broad band of radio frequencies.
Several studies talk about protocols and its characteristics, all the protocols developed for their

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

own specific needs and they are capable of filling these needs well.
We will mention some of them briefly in a table according to (Goldsmith, 2004) study.

Characteristic	Bluetooth	HomeRF
Operational Spectrum	2.402 - 2.480 GHz	2.404 - 2.478 GHz
Bandwidth	78 MHz	74 MHz
Modulation Type	FHSS (1600 Hops/sec), GFSK	FHSS (50 Hops/sec), 2-FSK, 4-FSK
Channel Access	Master-Slave Polling	CSMA/CA and TDMA
Data Rates	.721 Mbps Peak	.8, 1.8 Mbps
Data Traffic	PPP	TCP/IP
Range	Regular – 10 m High Power – 100 m	50 m
Error Robustness	1/3 rate FEC, 2/3 rate FEC, ARQ Type 1	CRC/ARQ Type I
Security	YES	YES
Communications Topology	Peer-to-Peer, Master-to-Slave	Peer-to-Peer, MS-to-BS
Vender Stability	Very Good	N/A
Device Scalability	Currently Very Low	Good
Data Scalability	Low	OK
Transmit Power	NA	100 mW
Energy Conservation	Yes	Directory Based
Capital Cost	Adapter: ~$30 Chipset: Under $4 in Bulk	N/A
Operational Cost	None	N/A

Wireless security has become just as important as the technology itself. This issue known in the media with much press on how easy it is to gain unauthorized access to a wireless network. It seems as if this attention has fallen on deaf ears as these networks are still incredibly in danger.
The absence of a physical connection between nodes makes the wireless links vulnerable to spy and information theft.

Insecure wireless user stations such as laptops create an even greater risk to the security of the enterprise network than rogue access points. The default configuration of these devices offer little security and can be easily misconfigured. Intruders can use any insecure wireless station as a launch pad to break in the network.

The basis for all WLAN security should start by understanding the environment in which your WLAN operates and its benifits.

We think about mobility and productivity as benefits of wireless, but that benefits put your information at risk.

We should pay attention on security alerts and set up a secure WLANs by implementing some practical actions.

(Khatod, 2004) implement five steps to protect the information assets, identify vulnerabilities and protect the network from wireless-specific attacks.

1. Discovery and improvement of Unauthorized WLANs and Vulnerabilities.
   it represent one of the biggest threats to enterprise network security by creating an open entry point to the enterprise network that bypasses all existing security measures including access points, soft access points (laptops acting as access points), user stations, wireless bar code scanners and printers.

   According to wireless security experts, discovery of unauthorized access points, stations and vulnerabilities is best accomplished with full monitoring of the WLAN.

2. Lock Down All Access Points and Devices
   The next step of WLAN security involves perimeter control for the WLAN. Each wireless equipped laptop should be secured by deploying a personal agent that can alert the enterprise and user of all security vulnerabilities and enforce conformance to enterprise policies. Organizations should deploy enterprise-class access points that offer advanced security and management capabilities.

3. Encryption and Authentication
   Encryption and authentication provide the core of security for WLANs. However ,fail-proo encryption and authentication standards have yet to be implemented.

4. Set and Enforce WLAN Policies
   WLANs needs a policy for usage and security.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

While policies will vary based on individual security and management requirements of each WLAN, a thorough policy and enforcement of the policy can protect an enterprise from unnecessary security breaches and performance degradation.

5. Intrusion Detection and Protection
   Security mangers rely on intrusion detection and protection to ensure that all components of WLANs are secure and protected from wireless threats and attacks.

To avoid the risks we should know it first, understanding how they work and using this information to avoid them as a solution for WLANs security.
A report from Internet Security Systems incorporation discuss some risks attacks against wireless technologies, they fall into seven basic categories:

1. Insertion attacks
2. Interception and unauthorized monitoring of wireless traffic
3. Jamming
4. Client-to-Client attacks
5. Brute force attacks against access point passwords
6. Encryption attacks
7. Misconfigurations

1- Insertion Attacks

Insertion attacks are based on deploying unauthorized devices or creating new wireless networks without going through security process and review (Bidgoli, 2006).

2- Interception and Monitoring of Wireless Traffic

As in wired networks, it is possible to intercept and monitor network traffic across a wireless LAN.
The attacker needs to be within range of an access point (approximately 300 feet for 802.11b) for this attack to work, The advantage for a wireless interception is that a wired attack requires the placement of a monitoring agent on a compromised system. All a wireless intruder needs is access to the network data stream.

3- Jamming

jamming can be a massive problem for WLANs. It is one of many exploits used compromise the wireless environment. It works by denying service to authorized users as legitimate traffic is jammed by the overwhelming frequencies of illegitimate traffic.

4- Client-to-Client Attacks

Two wireless clients can talk directly to each other, bypassing the access point. Users therefore need to defend clients not just against an external threat but also against each other.

5- Brute Force Attacks Against Access Point Passwords

Most access points use a single key or password that is shared with all connecting wireless clients. Brute force dictionary attacks attempt to compromise this key by methodically testing every possible password. The intruder gains access to the access point once the password is guessed.

6- Attacks against Encryption

802.11b standard uses an encryption system called WEP (Wired Equivalent Privacy). WEP has known weaknesses (see http://www.isaac.cs.berkeley.edu/isaac/wep-faq.html for more information), and these issues are not slated to be addressed before 2002. Not many tools are readily available for exploiting this issue, but sophisticated attackers can certainly build their own.

7- Misconfiguration

Many access points ship in an unsecured configuration in order to emphasize ease of use and rapid deployment. Unless administrators understand wireless security risks and properly configure each unit prior to deployment, these access points will remain at a high risk for attack or misuse.

Another report about Securing Wireless Local Area Networks suggests recommendations to protect wireless LAN network from attack, the following are some of them:

1. Educate employees about WLAN risks, and how to recognize an intrusion or suspicious behavior.
2. restrict unauthorized attachment of wireless access points (rogue access points).
3. Employ a third party managed security services company to constantly monitor the network security infrastructure for signs of an attack or unauthorized use.
4. Deploy strong for all of IT resources.
5. Ask users to connect only to known access points; masquerading access points are more likely in unregulated public spaces.
6. Deploy personal firewalls, anti-virus software and spyware blockers on all corporate PCs, particularly laptops and computers using the Windows operating system.
7. Actively and regularly scan for rogue access points and vulnerabilities on the corporate network, using available WLAN management tools.
8. Change default management passwords and, where possible, administrator account names, on WLAN access points.
9. Use strong security for other data resources such as laptop or desktop data files and e-mail messages and attachments.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

10. Avoid placing access points against exterior walls or windows.

11. Reduce the broadcast strength of WLAN access points, when possible, to keep it within the necessary area of coverage only.

12. Using of an Intrusion Detection System. This will provide your wireless network with early detection of common threats

Future works

Future work should focus on the following issues:

- Lack of method to detect a passive sniffer: An attacker usually first collects data traffic before launching an intrusion. This type of passive sniffing is quite dangerous, but there is nothing to do in this direction except to use the proper protection through encryption.

- To think about how to reduce and eliminate the risks attacks that can be happened on WLAN networks such as Man-in-the Middle attacks , Denial of Service (DoS) attacks and Identity theft (MAC spoofing)

- Authentication is the key: The most significant vulnerability of wireless LANs is the fact that, at the physical level, by definition they enable access to anyone, authorized or not, within a WLAN access point's radius of useful signal strength.

Conclusion

The future of wireless local-area networking is now, and it is the solution for communication problems in organizations or any place that need a wide spread of internet connection , interoperability became reality with the introduction of the standards and protocols and prices have dramatically decreased. These improvements are just a beginning.

Organizations who use WLANs networks can eliminate many of wireless LAN's security risks with careful education, planning, implementation, and management.

WLAN brings out not only advantages, but also some Specific security problems, although development of wireless standards and security protocols may enhance the WLAN security.

We know that hackers will never go away, so we bear the burden to provide the best 'locks' we can to protect our WLANs. Finally, whatever the outcome, wireless LANs will survive and are here to stay even if the technology has a new look and, or feel in coming years.

References

[1]Khatod, Anil, (2004). Fi**ve Steps To WLAN Security A Layered Approach.** AirDefense Inc. November 4, 2004 12:00 PM ET, http://www.computerworld.com/s/article/97178/Five_Steps_To_WLAN_Security_A_Layered_Approach

[2]Wireless LAN Security802.11b and Corporate Networks. An Technical White Paper, 2001, Internet Security Systems, Inc.

[3]Bidgoli, Hossein, (2006). **Handbook of Information Security, Threats, Vulnerabilities, Prevention, Detection, and Management.** Volume 3, Wily, 2006.

[4]Securing Wireless Local Area Networks. A VeriSign/Soltrus White Paper
2003 VeriSign, Inc. All rights reserved.

[5]Wireless Networking Basics, NETGEAR, Inc. October 2005, v1.0, October 2005

[6]Goldsmith, Colin, (2004). **Wireless Local Area Networking For Device Monitoring,** Master thesis, University of Rochester Rochester, New York

[7]Lansford, J., (2000). **HomeRFTM/SWAP: A Wireless Voice and Data System for the Home**. Intel Communications Architecture Labs, Hillsboro, Oregon, 2000

[8]O'Hara, B. & Petrick, A., (1999). **IEEE 802.11 Handbook: A Designer's Companion, Standards Information Network**, IEEE Press, New York, New York, 1999.

[9]The Wireless LAN Standard. Cisco Systems, 2000.

[10]802.11a: A Very-High-Speed, Highly Scalable Wireless LAN Standard., White Paper, 2002, www.proxim.com

[11]Wireless Networking Choices for the Broadband Internet Home., White Paper, 2001. www.homerf.org

[12]Wireless LAN Security. Symantec Corporation, 2002.

[13]Flickenger, Roger Weeks. (2005). **Wireless Hacks**, 2nd Edition, O'Reilly, 2005

[14]Clark, David, Pogran, Kenneth T. & Wed, David p. (1978). **An Introduction to Local Area Networks**. Proceedings of the IEEE, Vol. 66, 11, November 1978.

[15]Putman, Byron W.(2005). **WLAN Hands-On Analysis**. AuthorHouse, 2005.

[16]Aziz, Farhan Muhammad, (2003). **Implementation and Analysis of Wireless Local Area Networks for High-Mobility Telemetric.** Master

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Thesis submitted to the Faculty of Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

[17]Franklin, Tom, (2010). **Wireless Local Area Networks. TechLearn**, The Network Centre, Innovation Close, www.techlearn.ac.uk

[18]Holt, Keith, (2005). **Wireless LAN: Past, Present, and Future.** Intel Corporation.

[19]Negus, Kevin J., & Petrick, Al, (2009). **History of Wireless Local Area Networks (WLANs) in the Unlicensed Bands**. info, Vol. 11 Iss: 5, pp.36 - 56.

[20]Prem, Edward C., (2000). **Wireless Local Area Networks**. www.cis.ohio-state.edu/~jain/cis788-97/wireless_lans/index.htm.

[21]Chandramouli, Vijay, (2005). **Detailed Study on Wireless LAN Technologies.** http://crystal.uta.edu/~kumar/cse6392/termpapers/Vijay_paper.pdf, **2005**.

[22]Williams, Stuart K. (1999). **IrDA - Past, Present and Futur**e. Hewlett-Packard Company 2013.

[23]Kahn, Joseph M. & Barry, John R. (1997). **Wireless Infrared Communications**. Proceedings of the IEEE Vol. 85. NO. , February 1997.

[24]Gfeller, F. R. & Bapst,U. H., (1979). **Wireless in-house data communication via diffuse infrared radiation.** Proc. IEEE, vol. 67, pp. 1474–1486, Nov. 1979.

[25]Smyth, P. P., Eardley, P., Dalton, L. K., Wisely, T. D. R., McKee, P. & Wood, D., (1995). **Optical wireless: A prognosis.** in SPIE Proc. on [26]Wireless Data Transmission , vol. 2601, Philadelphia, PA, Oct. 23–25, 1995, pp. 212–225.

[27]Rancourt, .J. D., (1993). Safety of Laser Products. Int. Electrotech. Commission, CEI/IEC825-1: Optical Thin Films. New York: Macmillan.

[28]Carruthers, Jerrey B., (2002). Wireless Infrared Communications. Wiley Encyclopedia of Telecommunications

[29]http://www.irda.org/

[30]Cisco Systems, Inc. (2000).

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Ambient noise Coherence properties detection for various Hydrophone Spacing

V.G.Sivaumar
Department of ECE
Sathyabama University
Chennai, India

Dr.V.Rajendran
Department of Physics/Ece
SSN College of Engineering
Chennai, India

***Abstract*—Ambient noise is a complex and important phenomenon which greatly affects the listening capacity of instruments in underwater environment. The ambient noise in sea is the overall combination of wind speed, wave speed, wave height, barometric pressure, dew point, temperature, marine life, shipping traffic and seismic activities. The present work concentrates on coherence with various hydrophone spacing. Under water ambient noise analysis is essential to enhance the Signal to Noise Ratio (SNR) of acoustic based underwater instruments. This paper investigates the effect of noise spectrum over a different hydrophone spacing and the signal coherence with hydrophone spacing is examined in the Bay of Bengal Sea region.**

Keywords-component; Ambient noise; Noise Level; Wind speed; Coherenc.

I. INTRODUCTION

The noise environment of the ocean is an important aspect of habitat for marine mammals and other organisms, and introduction of human-generated sounds may result in auditory, physiological, or behavioral impacts (National Research Council, 2003 and 2005). Study of noise is very important for the design and development of underwater acoustic instruments such as Sonar, echo sounder etc., and also for acoustic communications. Ambient noise is a limiting factor in the performance of underwater acoustic detection and communication systems in shallow water [5]. Ocean ambient noise is an inherent characteristic of the medium having no specific point source. It is the residual noise background in the absence of individual identifiable sources that may be considered as the natural noise environment for hydrophone sensors. It comprises of number of components that contribute to the Noise Level (NL) in varying degrees depending on the location of measurements [Urick R., 1983]. As performance of any underwater communication systems

Rely on the signal to noise ratio, the information on ambient noise field is required since then background noise masks the signals from the systems. Hence, measurements, analysis and characterization of ambient noise at any location in the ocean is necessary. Sea surface generated noise is the prevailing noise of the ocean and its importance was identified in the earliest studies of ambient noise. It was originally considered to be a function of sea state, but later studies found that the noise correlated better with wind speed [1-2] and it has since been known as "wind-dependent noise". Ambient noise is the sustained unwanted background noise prevailing at any location. This masks the signals from underwater acoustic instruments. So the detection of background noise is essential to enhance the Signal to Noise Ratio (SNR) of acoustic based underwater instruments. A direct connection between wind force and the level of ambient noise is observed for a frequency range of 500 Hz to 25 kHz. Noise level spectrum is summarized in [3]. Knudsen spectra [4] show the strong dependence of spectral power level with wind speed and sea states. The properties of Noises in shallow waters, along with coherence and vertical directionality is seen to exhibit site specific characteristics where speed of sound and bottom properties vary substantially with location. Numerical models for shallow water environment have been developed [7,8,9] in which the spatial coherence of noise field is computed and they conclude that the form of the noise field depend on whether the bottom is a low loss or a high loss boundary. Vertical coherence of noise off the Scotian shelf have been studied by Desharnais et al. [11] and inferred that at frequencies dominated by sea surface noise, coherence is employed for understanding bottom properties of the region. In this paper, the power spectrum is estimated for the Bay of Bengal Sea region. Finally the ambient noise coherence in shallow waters of Bay of Bengal is estimated.

II. METHODOLOGY

A. Data Collection

The ambient noise data is collected by placing five hydrophones vertically at a depth of 12 meter in the bay of Bengal region. The collected data's are recorded using the data acquisition system. The ocean depth in the Bay of Bengal sea site was 12 meters and the sampling frequency used while taking measurement was 200 KHz. The wind speed in this region ranges from 0.95 m/s to 6.56 m/s. The data collection process were taken for the period of one week with sea state of moderate wind speed around 0.95m/sec to 6met/Sec. The following figure shows the sample method of data collection setup.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

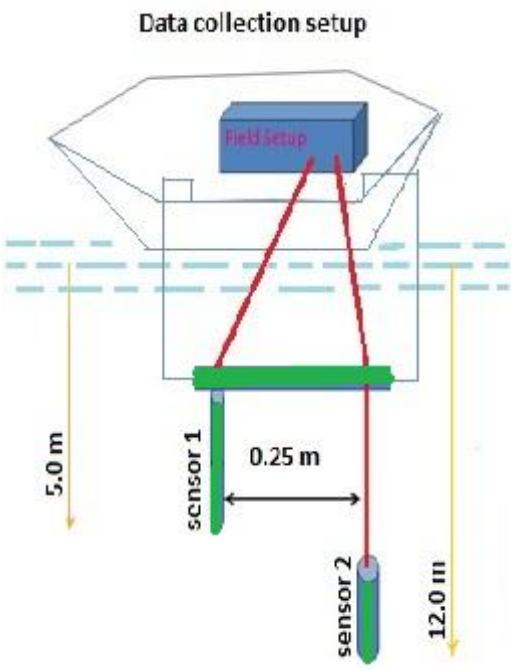


Fig1.A sample model of Ambient noise setup for data collection

Power Spectral Density Estimation

The Spectral analysis is key to understanding signal characteristics, and it can be applied across all signal types, including radar signals, audio signals, seismic data, financial stock data, and biomedical signals. The goal of spectral estimation is to describe the distribution (over frequency) of the power contained in a signal, based on a finite set of data. Estimation of power spectra is useful in the detection of signals buried in wideband noise. Power spectral density refers to the amount of power per unit of frequency as function of the frequency. The power spectral density describes how the power (or variance) of a time series is distributed with frequency. By knowing the power spectral density and system bandwidth, the total power can be calculated. The power spectral density (PSD) of a stationary random process x_n is mathematically related to the autocorrelation sequence by the discrete-time Fourier transform. In terms of normalized frequency, this is given by

$$P_{xx}(w) = \frac{1}{2\pi}\sum_{m=-\infty}^{\infty} R_{xx}(m)\, e^{-j\omega m} \quad \text{----------(1)}$$

The average power of the sequence x_n over the entire Nyquist interval is represented by

$$R_{xx}(0) = \int_{-\pi}^{\pi} P_{xx}(\omega)\, d\omega = \int_{-f_x/2}^{f_x/2} P_{xx}(f)\, df \quad \text{---------(2)}$$

Where, $P_{xx}(\omega)$ represents the power content of a signal in an infinitesimal frequency band, which is why it is called the psd.

Welch's Method :

Welch's method (also called the averaged modified periodogram *method*) for estimating power spectra is carried out by dividing the time signal into successive blocks, forming the periodogram for each block, and averaging. Denote the m^{th} windowed, zero-padded frame from the signal *x* by

$$x_m(n) \triangleq w(n)x(n+mR), \qquad n = 0,1,\ldots M-1, m = 0,1,..,K-1$$

Where, *R* is defined as the window *hop size*, and let *K* denote the number of available frames. Then the periodogram of the m^{th} block is given by

$$P_{x_m,M}(w_k) = \frac{1}{M}\left|FFT_{N,k}(x_m)\right|^2 \triangleq \frac{1}{M}\left|\sum_{n=0}^{N-1} x_m(n)\, e^{-j2\pi nk/N}\right|^2 \quad \text{----(3)}$$

as before, and the Welch estimate of the power spectral density is given by

$$\hat{S}_x^W(w_k) \triangleq \frac{1}{K}\sum_{m=0}^{K-1} P_{x_m,M}(w_k) \quad \text{----------(4)}$$

In other words, it's just an average of periodograms across time. When *w(n)* is the rectangular window, the periodograms are formed from non-overlapping successive blocks of data.

Coherence:

Coherence determines the similarity between the signal measured in two hydrophones. The coherence equation is given by

$$r(f)_{xy} = \frac{|s_{xy}(f)|^2}{s_{xx}(f)\, s_{yy}(f)} \quad \text{----------(5)}$$

Where, S_{xx} , S_{yy} are the auto spectral density functions and the mathematical Equations are,

$$S_{xx} = \int_{-\infty}^{\infty} R_{xx}(\tau)\, e^{-j2\pi f\tau}\, d\tau \quad \text{----------(6)}$$

Identify applicable sponsor/s here. *(sponsors)*

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

$$S_{yy} = \int_{-\infty}^{\infty} R_{yy}(\tau)\, e^{-j2\pi f\tau}\, d\tau$$ ----------(7)

Then, S_{xy} denotes the cross spectral density is given as

$$S_{xy} = \int_{-\infty}^{\infty} R_{xy}(\tau)\, e^{-j2\pi f\tau}\, d\tau$$ ----------(8)

III. Results And Discussion

A. Spectral Analysis

The noise spectrum has been plotted for the wind speeds 5.34m/s in the Bay of Bengal sea data. The fig2 shows that the power spectrum for First hydrophone in hydrophone array at hydrophone depth of 12m. It can be noted form the fig that at 48.21Hz Hz the noise spectrum level is 133.8dB for a wind speed of 5.34 m/s and it is 131.9 dB for same frequency, wind speed. It is clearly evident that due the spacing between the hydrophones (0.5metter) there is an small variation in the noise level. By comparing both figures that is fig2 and fig3 the noise level varies only for the low frequencies ranges from 0 to 5KHz.In the high frequency ranges the noise levels are in almost stable condition for both hydrophones. The figure.4 shows that wind speed Vs Nose level for various frequency ranges. It is clearly proved that the noise level increases with increasing of wind speed

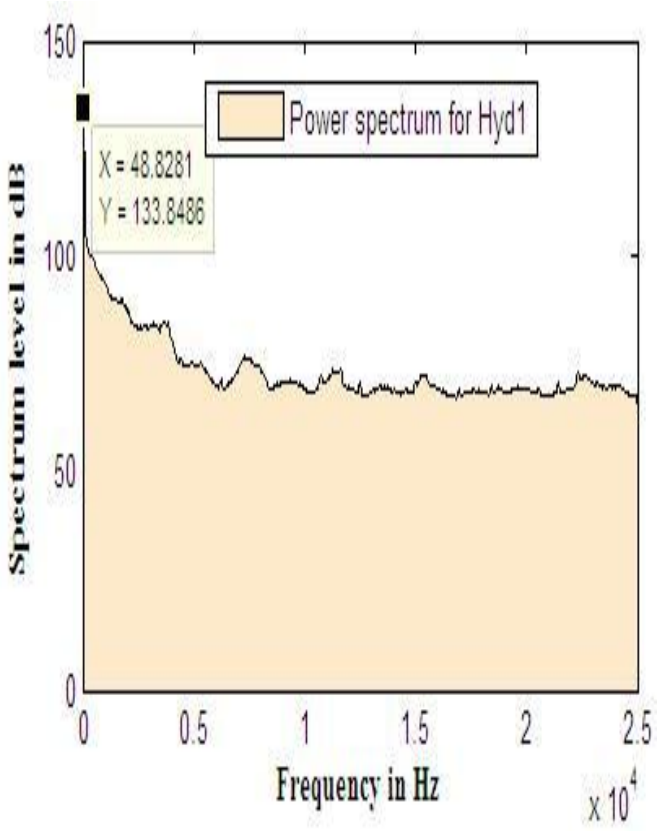


Fig2. Power spectrum Estimation in shallow water region of Bay of Bengal.

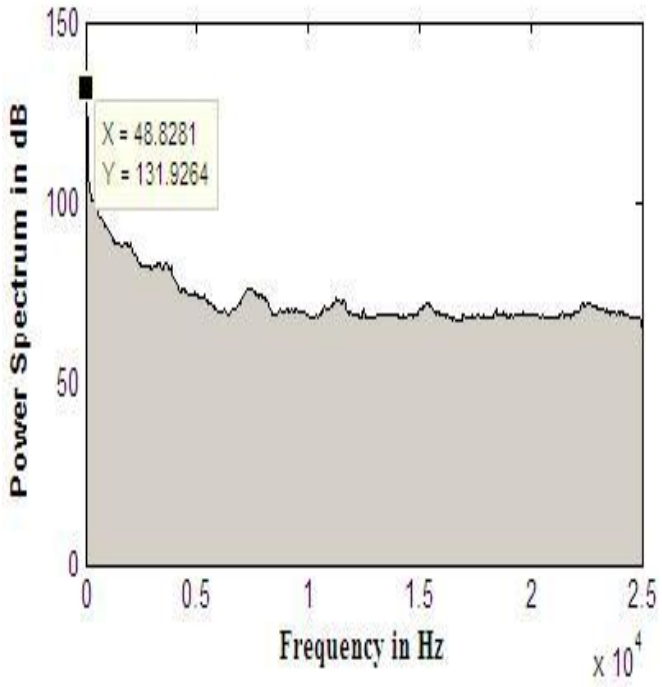


Fig.3 (a) Power spectrum estimation for Hyd2- in shallow water region of Bay of Bengal .

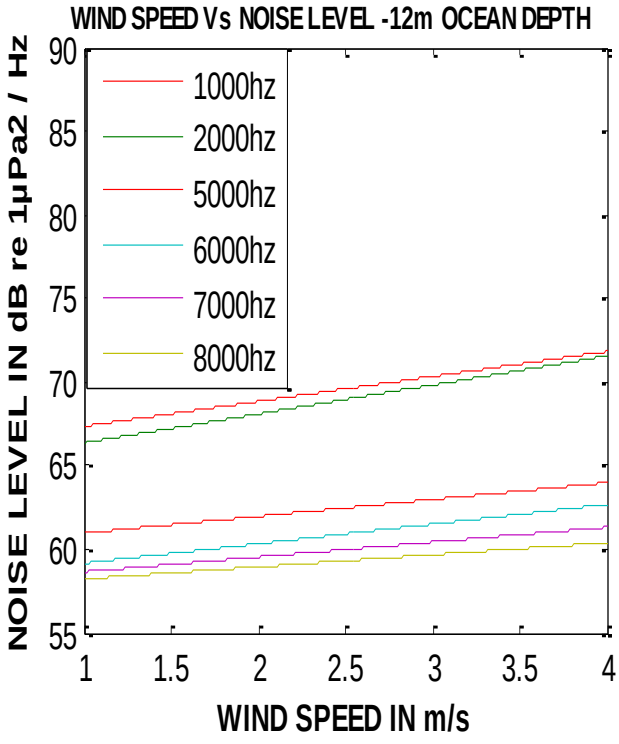


Fig.4 Noise Level Vs wind speed for different frequency

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

B. Coherence

Coherence is a normalized measurement. So that the vertical coherence of ambient noise field is expected to be stable over a wide range of environmental conditions. The vertical coherence in Bay of Bengal region is shown below. The fig 5, fig 6 shows the real coherence and imaginary coherence at 5.34 m/s wind speed. The top most graphs in fig5 shows the real coherence between the hydrophone of 1&2. The middle figure shows real coherence of 1&3 and bottom shows the real coherence of 1&4.In this research the properties of coherence is estimated that is due to the variation in the hydrophone spacing the coherence also varies.

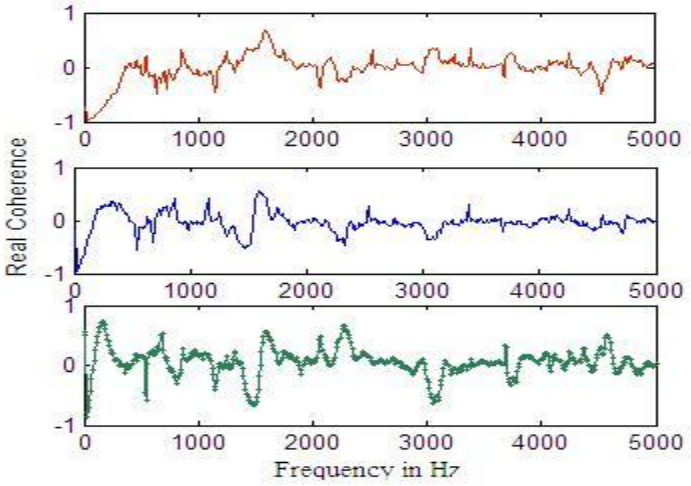


Fig.5.Comparision of Real Coherence for various hydrophone Spacing

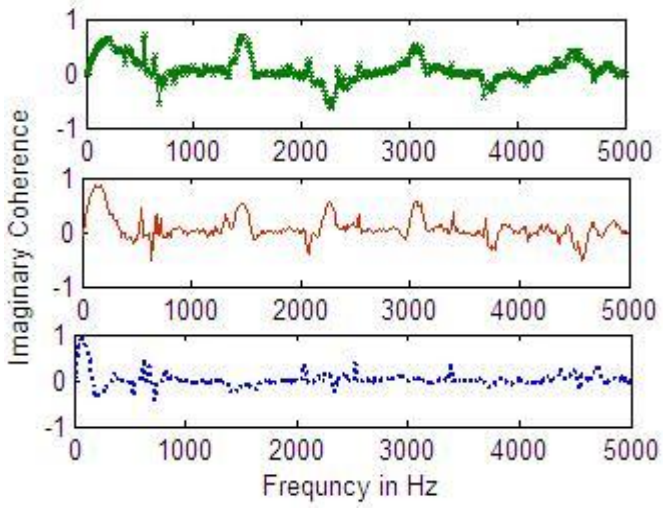


Fig.6.Comparision of imaginary Coherence for various hydrophones Spacing

The figure5 intimate that once hydrophone spacing increases the coherence level also increases. The topmost graph in fig5 shows the coherence level for hydrophone spacing is 0.5mts and the second graph is 1mt and the final one is 1.5mts.From the figures it is clearly understand that if spacing increases the noise coherence level also increases. One of these properties will be used to estimate the properties of sea bed. The figure 6 shows how the imaginary coherence level varies with different hydrophone spacing.

IV. Conclusion

Underwater ambient noise is a very complex and critical one to predict for different sea state. The data collected by five element of the vertical line array for a period of one week. In this research we obtained the power spectrum of ambient noise signal for various hydrophone spacing and also we predicted the coherence of hydrophone pair for different hydrophone spacing. From the out put it is concluded that for low-frequency signal the power spectrum can be predicted but for higher frequencies we can't accurately predict the noise spectrum. It is clearly evident that the coherence level varies with different hydrophone spacing in an vertical array. Several aspects will be focused in future works. We will focus on two things mainly: 1) the computation and the analysis of the coherence function and power spectrum with varies hydrophone spacing for this particular Bay of Bengal sea region, and 2) From this research how this coherence properties will be used to estimate the seabed properties various sea state conditions.

References

[1] Tan Soo Peing, Koay Teong Beng, P.Venugopalan, Mandar A Chitre and John R.Potter. "Development of a Shallow Water Ambient noise Database", Acoustic Research Laboratory, Tropical Marine Science Institute, National University of Singapore.

[2] Urick RJ. Ambient noise in the sea. Peninsula Publishing; 1984.

[3] D.H. Cato, S.Tavener. "Ambient sea noise dependence in local regional and geostrophic wind speeds: Implications for forecasting noise". *Applied acoustics*, volume 51, Issue 3,1997, Pages 317-338.

[4] D.H. Cato, S.Tavener. "Wind dependence of ambient noise in shallow water of Bay of Bengal". *Applied acoustics*, volume 69, Pages 1294-1298.

[5] Wenz G.M. (1962), 'Acoustic ambient noise in the ocean: Spectra and sources', Journal of Acoustic Society of America, Vol. 34, pp. 1936-1956.

[6] Knudsen V.O., Alford R.S. and Emling J.W (1948), 'Underwater ambient noise', Journal of Marine Research, Vol. 7, pp 410-429.

[7] Buckingham, M. J. (1980). A theoretical model of ambient noise in a low loss, shallow water channel. J. Acoust. Soc. Am., 67,1186-1192.

[8] C H Harrison (1995), Formulas for ambient noise level and coherence. J. Acoust. Soc. Am., 99(4), 2055-2066.

[9] Desharnais.F, MacDonald .B.R and Mah K.J. (1998). Vertical Noise Coherence measurements in shallow water using lagrangian drifters, Defence Research Establishment Atlantic, Report 507466.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[10] Sanjana M.C, G. Latha and Rajendran. V(2009). Vertical Coherence of ambient noise in shallow waters of Bay of Bengal.

[11] M. J. Buckingham, ‘‘A theoretical model of ambient noise in a low-loss,shallow water channel,’’ J. Acoust. Soc. Am. 67, 1186–1192 ~1980!.

AUTHORS PROFILE

V.G.Sivakumar was born in Tamilnadu, India on July 1st, 1972.He received his B.E. (Electronics and Communication Egg.) degree from Bangalore University, India, in 1998, M.E.(Applied Electronics) degree from Sathyabama University, India ,in 2004. He was worked as an service Engineer in Hi-Tech Software centre (1998-2000) and worked as an junior software Engineer in Kaashyap Radiant system from the year 2000-2002.Presently he is working as an Assistant professor in Sathyabama University Chennai, India. and doing his research in underwater acoustic signal processing.

Dr.V.Rajendran, Graduated from MK University, **C**ompleted his M.Tech.from IISC Bangalore and received doctorate from Chiba University, Japan in 1993. He has been working in different institutions like Indian Institute of Science (IISc), Bangalore, Indian Institute of Technology (IIT), Madras and NIOT Chennai. He received MONBUSHO Fellowship Award of Japanese Government and Distinguished Scientist Award ’07 from Jaya Engineering College. He has also been Elected Member twice as Vice Chairman - Asia of Executive Board of Data Buoy Co-operation Panel (DBCP) of Inter- Governmental Oceanographic Commission (IOC) / World Meteorological Organization (WMO) of UNSCO, in October 2008 and September 2009.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

ADAPTIVE IRIS LOCALIZATION AND RECOGNITION: MODIFICATION ON DAUGMAN'S ALGORITHM

Marwan AL-abed Abu-zanona *
Department of Computer Science
Imam Muhammad Ibn Saud Islamic University, KSA

Bassam M. El-Zaghmouri
Department of Computer Information Systems
Jerash University, Jordan

Abstract— the use of biometric information has been widely known for both people identification and security application. It is common knowledge that each person can be identified by the unique characteristics of one or more of biometric features. One most unique and identifiable biometric characteristics is the iris, wherever the second is the voice, and the third is finger print. This research attempts to apply iris recognition techniques based on the technology invented by Dr. John G. Daugman, an attempt of implementing a build an end user application. Iris Recognition is expected to play a major role in a wide range of applications in which a person's identity must be established or confirmed in high reliability and high privacy, Including access controls, authorizations, ID detection, etc. This research depends on standard iris images was token from CASIA database. The most efficient computer language for simulation and technical computing (MATLAB) will be used to make the problem statement and result in addition to mathematical and AI modelling more easier and reliable.

Keywords— Image Processing; Iris; Localization; Biometrics; Gradient

I. INTRODUCTION

Human Identification / Verification are an ancient goal of the humanity. Hence the technology and its services have developed in the modern world, human activities and transactions have increased in which rapid and reliable personal identification is required. Many examples include computer login control, passport control, bank automatic teller machines and other transactions authorization, access control, and security systems in general. All such identification efforts have the common goals of speed, reliability and automation [1].

Recorded voiceprints are susceptible to changes in many parameters affects the person's voice, systematic factors, non-systematic effects, and they can be counterfeited. Fingerprints or handprints require physical contact, and they also hard to implement and usage [7].

On the other hand, human iris print is an internal organ of the eye and had a special protection against the external environment. It is easily visible from within one meter long distance. This makes it perfect biometric information for an identification/verification system with the ease of speed, reliability and automation [2].

This research, experiment, implements, and also, looks into the theory of the Iris Recognition System. This related to the field of personal identification / verification and more specifically to the field of automated identification / verification of humans by biometric information.

II. IRIS BIOMETRICS

Any biometrics should have the specific attributes. The major one is the (DOF) degree-of-freedom of variance in the specified index related to the human population. This determines the uniqueness; its immutability over time and its immunity to intervention. The second

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

attribute is the computational prospects for efficiently encoding and reliably recognizing. In the entire human population, no two irises are alike in their mathematical detail, even among identical (monozygotic) twins. The probability that two irises could produce exactly the same Iris Code is approximately 1 in 1078. (The population of the earth is around 1010) [14].

Possibility of using the iris of the eye as a kind of personal identification / verification optical like Fingerprint was first suggested by ophthalmologists who noted from clinical experience that every iris had a highly detailed and unique texture, which remained unchanged in clinical photographs spanning decades. The iris is composed of elastic connective tissue, the trabecular meshwork, whose prenatal morphogenesis is completed during the 8th month of gestation. It consists of pectinate ligaments adhering into a tangled mesh revealing striations, ciliary processes, crypts, rings, furrows, a corona, sometimes freckles, vasculature, and other features [2], [13], [14].

The color of the iris is usually changed by blanket of chromatophore cells during the first year of life, but trabecular pattern itself is stable throughout the lifespan according to the available clinical evidences. Properties that enhance its suitability for use in automatic identification include: its inherent isolation and protection from the external environment, being an internal organ of the eye, behind the cornea and the aqueous humor; the impossibility of surgically modifying it without unacceptable risk to vision and its physiological response to light, which provides a natural test against artifice [13][14].

The iris is shared with fingerprints in the property of random morphogenesis of its minutiae. The iris texture is stochastic or possibly chaotic. That is because of there is no genetic penetrance in the expression of this organ beyond its anatomical form, physiology, color and general appearance [13][14].

Because of the detailed morphogenesis of the iris depends the embryonic mesoderm's initial conditions from which it develops the phenotypic expression even of two irises have the same genetic genotype, they must have uncorrelated minutiae. Thus, the uniqueness of specified fingerprint parallels the iris uniqueness, common genotype or not. But more advantages in particular can be extracted from this [12].

III. METHODOLOGY

After we acquire the image using camera the first stage of iris recognition is to isolate the actual iris region in a digital eye image. The iris region, shown in Figure 1.2, can be approximated by two circles, one for the iris/sclera boundary and another, interior to the first, for the iris/pupil boundary. The eyelids and eyelashes normally occlude the upper and lower parts of the iris region. Also, specular reflections can occur within the iris region corrupting the iris pattern. A technique is required to isolate and exclude these artefacts as well as locating the circular iris region [1].

The accuracy of iris segmentation depends on the imaging quality, and the pre-processing of the eye image. Images in the CASIA (the most famous, most used, and what we work on in this research) iris database do not contain specular reflections due to the use of near infra-red light for illumination. However, the images in the LEI database contain these specular reflections, which are caused by imaging under natural light. Also, persons with darkly pigmented irises will present very low contrast between the pupil and iris region if imaged under natural light, making segmentation more difficult. The segmentation stage is critical to

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

the success of an iris recognition system, since data that is falsely represented as iris pattern data will corrupt the biometric templates generated, resulting in poor recognition rates [3].

A broad set of image processing operations that affects the image depending on shapes is so called MORPHOLOGY. Morphological operation applies a structure element to the image, the output resulted image will be the same size. Each pixel' value in the output image is based on a comparison of the corresponding pixel in the input image with its neighbors. By choosing the size and shape of the neighborhood, you can construct a morphological operation that is sensitive to specific shapes in the input image [15].

The basic morphological operations is "Opening in" and "Closing". Morphological opening and closing changes the definition of pixel set depending on the neighborhood pixels and structure element. And the most basic operations in morphology are dilation and erosion. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. In the morphological dilation and erosion operations, the state of any given pixel in the output image is determined by applying a rule to the corresponding pixel and its neighbors in the input image. The rule used to process the pixels defines the operation as dilation or erosion. This table lists the rules for both dilation and erosion [15].

The structure elements used in this research is "Liner" and "Disk". Structure elements are the basic block of any morphological operation. Equation 1 represents "Linear" structure element, while the equation 2 represents a "Disk" structure element [15].

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \qquad (1)$$

$$\begin{matrix} 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \end{matrix} \qquad (2)$$

Controlling the structural elements will generate a different images and different results of processing. The programmer matter is to determine and design the best structure element and the best parameter of that [15].

IV. RADON TRANSFORMAITON

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

A standard mathematical model to represent the projection of geometric 2-D object into one dimension is implemented in computer vision algorithms in the name “Radon Transform”. That transform is usually used to determine the parameters of simple geometric objects by the means of projection, such as lines and circles, inside the image. Radon transform is often so called “Hough Transform” in image processing. It usually employed to detect the tangents and centre coordinates of the circular or curved regions. It’s very efficient in line detection [12].

In this research, an automatic segmentation algorithm based on Hough transform will build and tested. After building edge map in the eye image the Hough transform will be applied to specify parameters of circles passing through each edge point. These parameters are the centre coordinates xc and yc, and the radius r, which are able to define any circle [17].

The lines in the image is determined by the maximum point of projection in radon space in the Hough space and the corresponding radius and centre coordinates of the circle will be best defined by the edge points of the tangential line. Approximating the upper and lower eyelids with parabolic arcs will be available using radon transform based line detection [17].

According to the experimental results, the error in determining the iris coordinated and border should be substituted in a specific mathematical computational algorithm. The gradient is defines the variance between a set of mathematical data. Hence, the gradient can be implemented to get the maximum variance in between the iris and cornea [12].

The eyelid edge map will corrupt the circular iris boundary edge map if using all gradient data. Considering the vertical gradients only for locating the iris boundary will efficiently reduce influence of the eyelids after performing the radon transform. Horizontal gradient will be very useful in locating the iris boundary. Not only does this make circle localization more accurate, it also makes it more efficient, since there are less edge points to cast votes in the Hough space [17].

Two-dimensional gradient mathematical equation is described in equation 3 [12].

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - \nabla f(x) \cdot h\|}{\|h\|} = 0 \tag{3}$$

V. SYSTEM DIAGRAM

This proposed systems works in two modes; the first is enrollment mode, and the second is identification mode. In the first mode, the templates will be taking and the iris code stored in the database. The second is the running mode in normal condition to get identification.

Figure 1 illustrates the program flow.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

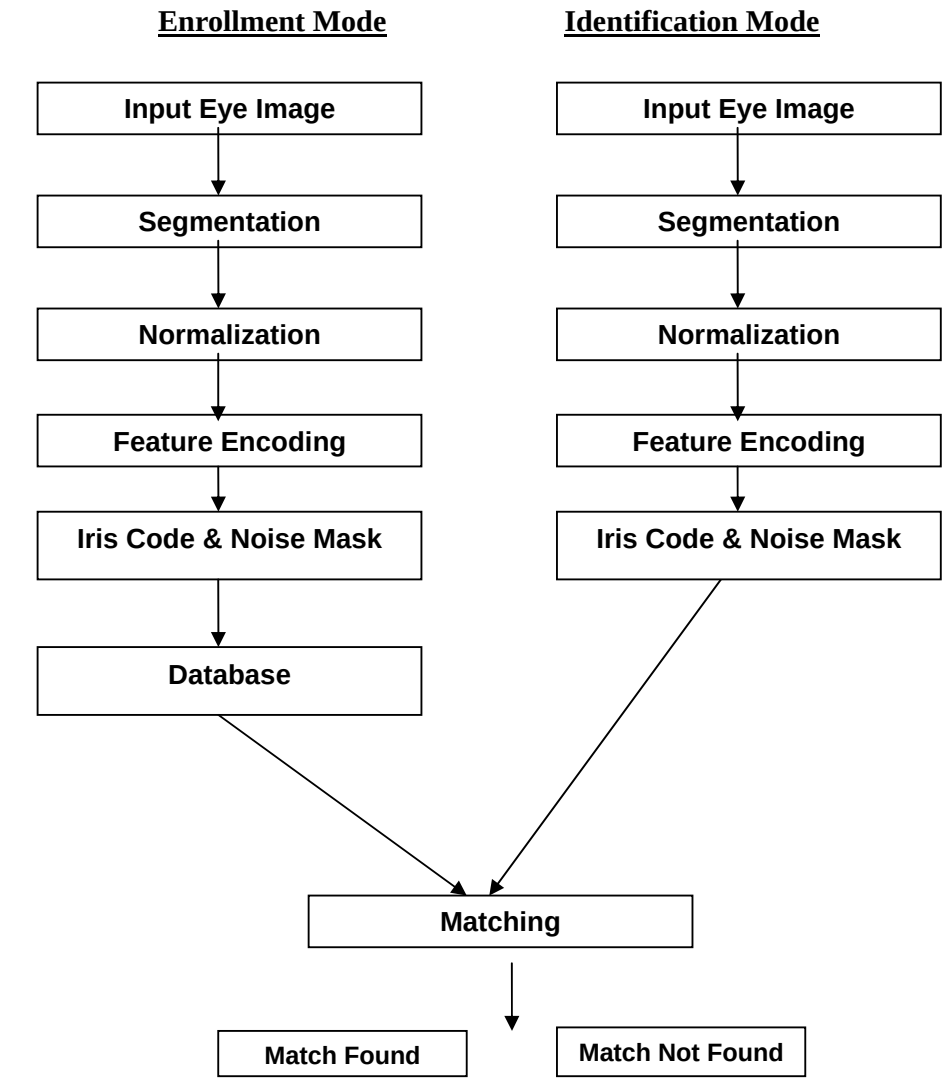


Fig. 1: Proposed System Flow Chart

VI. IRIS CODE

Feature encoding was implemented by convolving the normalized iris pattern with 1D Log-Gabor wavelets. The 2D normalized pattern is broken up into a number of -D signals, and then these 1D signals are convolved with 1D Gabor wavelets. The rows of the 2D normalized pattern are taken as the 1D signal; each row corresponds to a circular ring on the iris region. The angular direction is taken rather than the radial one, which corresponds to columns of the normalized pattern, since maximum independence occurs in the angular direction [3].

The iris intensity values at specific known noise areas in the normalized pattern are set to the mean intensity of neighborhood pixels to remove the influence of noise in the filter's output. The filtering output is then phase quantized to four levels using the Daugman method, with each filter producing two bits of data for each phase. The phase output quantization should be chosen to be a grey-level code, thus, when sliding between two quadrants, only one bit will change. This is minimizing the number of bits disagreeing, and thus will provide

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

more accurate recognition. The feature encoding process is illustrated in Figure 2.The result code is so called “Iris Code” [3].

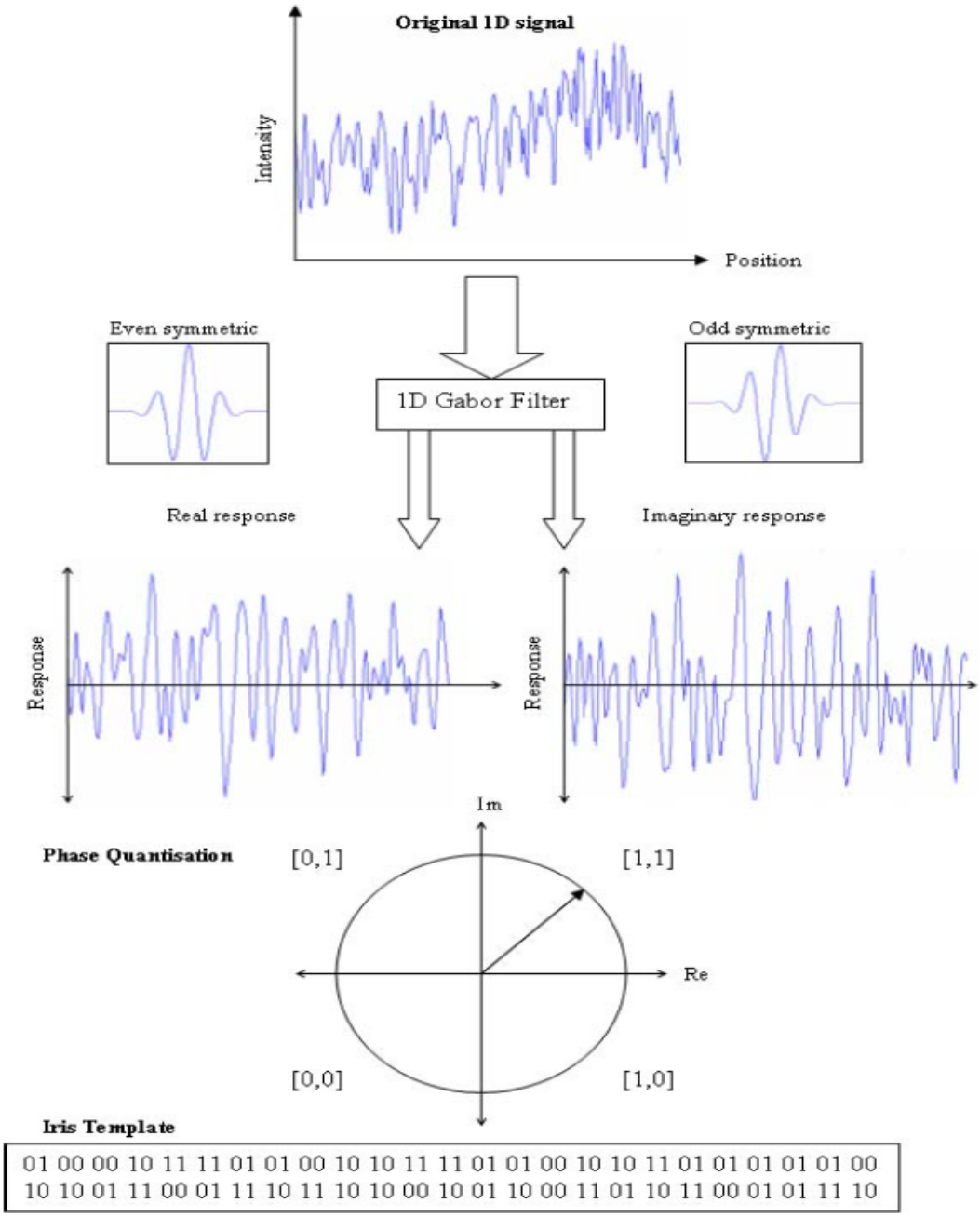


Fig. 2: Iris Code

A bitwise template is produced by the encoding process. This template is what so called “Iris Code” containing number bits carrying the information of the iris, and the noise mask which corresponds to corrupt areas within the iris pattern, and marks bits in the template as corrupt. The phase information is meaningless at the regions of zero amplitude, so, the noise mask will also mark these regions. The total number of bits in the iris template will be the radial

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

resolution times the angular resolution, times 2, times the number of used filters. In this technique, the number of used filters and their centre frequencies, and the parameters of the modulating Gaussian function are responsible to achieve the best recognition rate. Figure 3 shows the iris code mask [1].

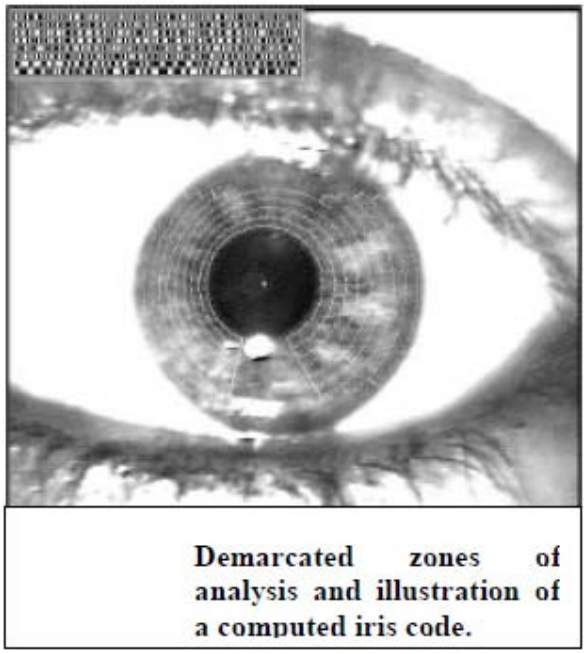


Fig. 3: More Illustration on Iris Code

VII. MATCHING

Daugman patented algorithm is concerns on matching the iris depending on Hamming distance represents the measure of how many bits are the same between two patterns of bits. The Hamming distance should be used between two bit patterns to generate a decision that can be whether the two patterns were generated from different irises or from the same one. For example the comparison between the two bit patterns X and Y, the equation Hamming distance, HD, is defined as the sum of disagreeing bits over N, the total number of bits in the bit pattern. HD is described in equation 4.

The HD for two codes generated from the same iris will be less than 0.3, and it larger than 0.3, the matching will get fail result (not match). Figure 4 shows examples of HD on different patterns [1].

$$HD = \frac{1}{N} \sum_{j=1}^{N} X_j (XOR) Y_j \quad (4)$$

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

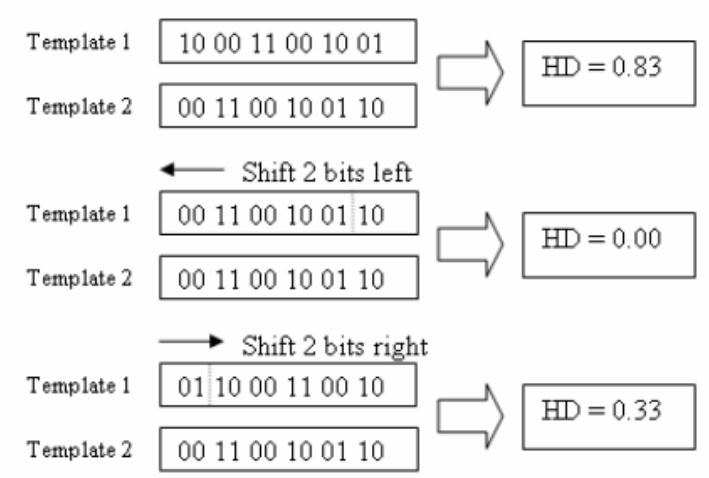


Fig. 4: Examples of Applying HD

VIII. RESULTS

The following figure illustrates the program flow from starting the iris image, passing through morphological preprocessing, Hough transformation, iris localization, and pattern isolation.

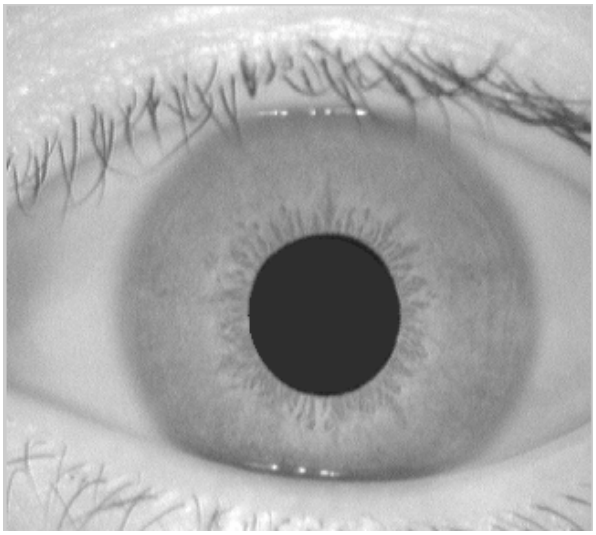


Fig. 5: Original Iris Image

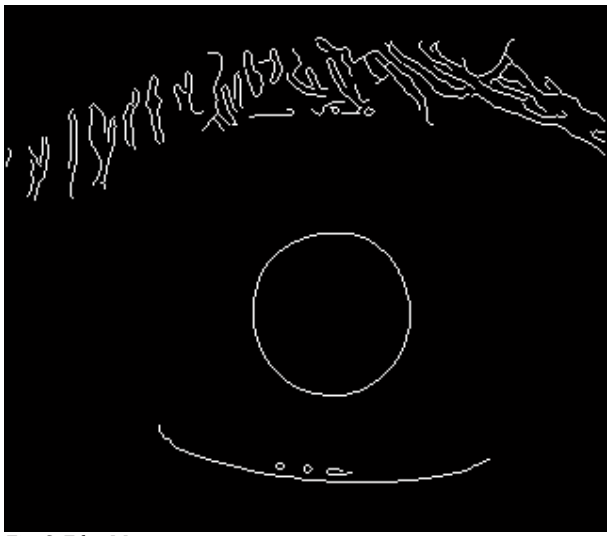


Fig. 6: Edge Map

The original image should be gray or converted to gray (see section 2.1). Sample iris is shown in figure 5, it relates to CASIA data base.

First preprocessing is the finding edge map using first derivative (Laplacian), edge map will enable to localize the pupil and determining the center of pupil. It could be used in gradient after some steps. Figure 6 shows the first edge map of the iris.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

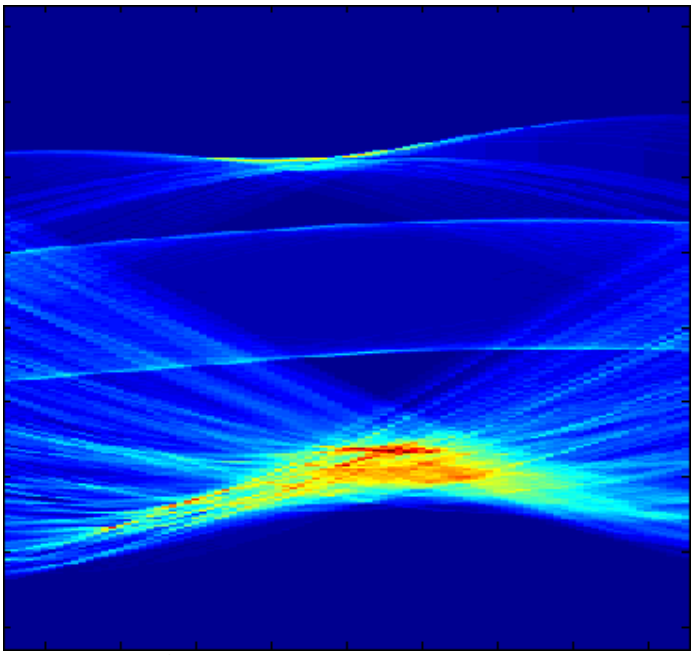


Fig. 7: Radon Transformation Diagram

The projection representation of the radon transform is shown in figure 7. It's clear from this representation that the pupil can't be isolated from this projection, because of the concentrated effect of eyelashes.

Reconstructing the image after transformation will result the image in figure 8, it's clear that the eyelashes is easier to localize after reconstruction. So, the combination of the Hough transform with morphology and computational mathematics would result best localizing of the iris.

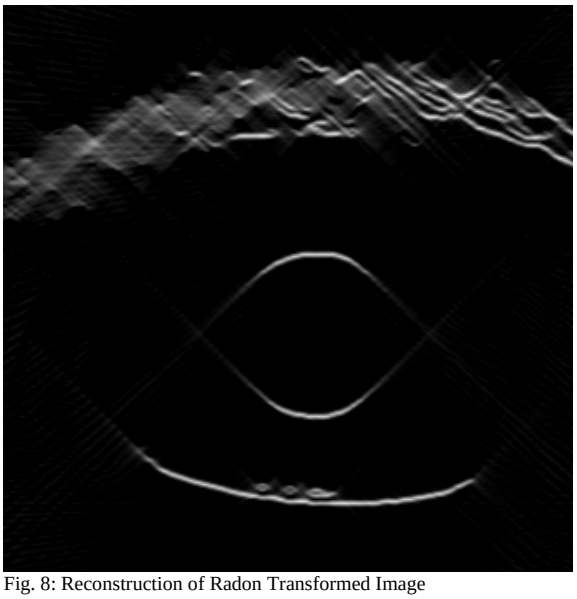


Fig. 8: Reconstruction of Radon Transformed Image

Fig. 9: Dilated Image

Figure 9 shows the dilation morphology of the image. This improves in connecting the objects which has some cutting or some erosion.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

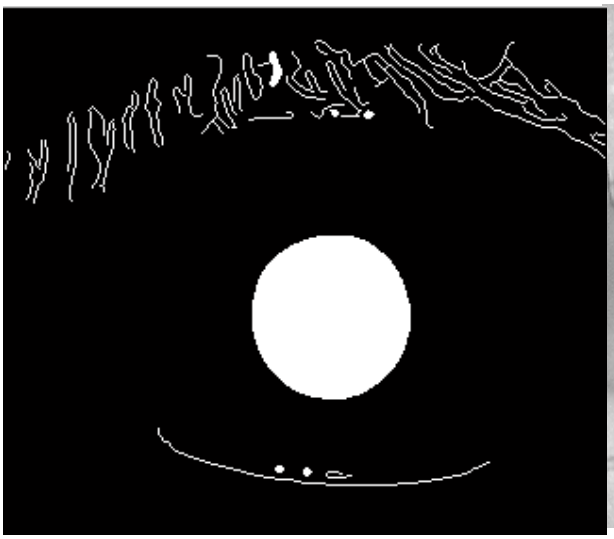


Fig. 10: Localizing the Pupil as Solid Object

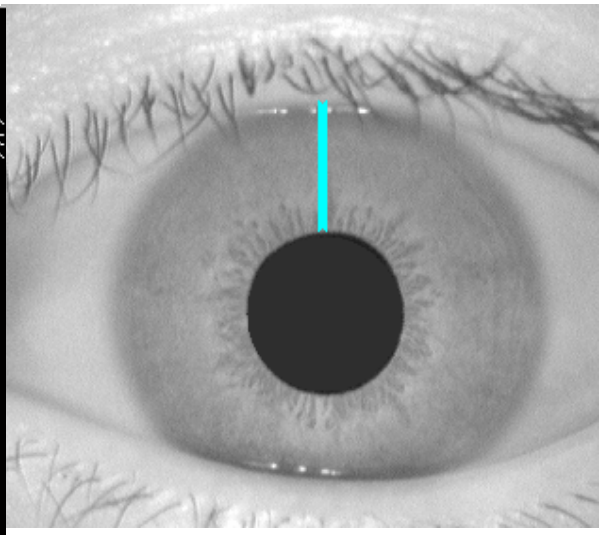


Fig. 11: Vertical Gradient

Now, to determine the centroid of the pupil, filling of the closed objects in edge map will close the pupil. After that, the region properties of it would be easy to calculate. The other processes will use the centroid of the pupil. Figure 11 shows the vertical gradient path in mathematical calculation.

Attempting to localize the iris is done using Hough transformation and gradient to substitute its error. First, using the Hough transformation to determine the eyelashes, the error in this way should be substituted. Figure 12 illustrated the iris localization using Hough transform in red color.

The blue circle in figure 12 illustrates the starting of gradient calculations. This circle limit was been found using the morphology. The final step is calculating the gradient between the pupil limit and the end of the eye. The contribution in such way is minimizing the error or Hough transform. The green circle in Figure 12 ensures the good performance of proposed gradient method.

The next step after localizing the iris is isolating it. Mathematical circle equation is used in geometry to make every pixel out of the iris circle returns zero, as Figure 13.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

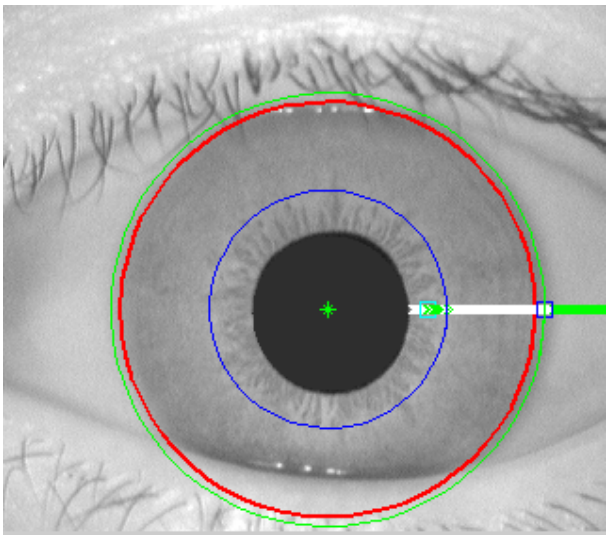


Fig. 12: Localizing the Iris

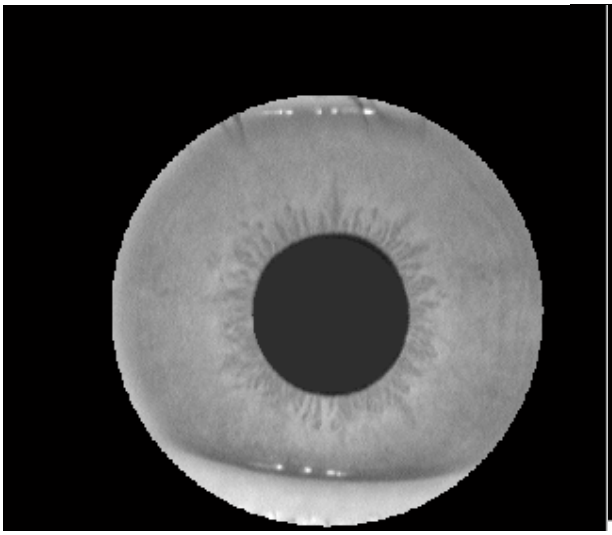


Fig. 13: Detecting the Iris

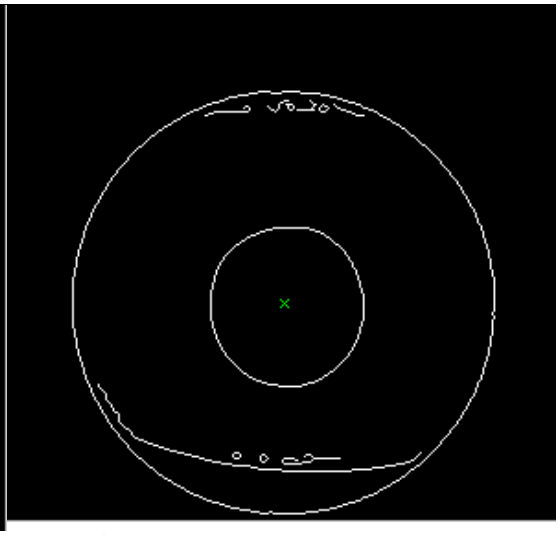


Fig. 14: Edge of the Detected Iris Circle

After the past isolation of the iris, a part of eyelids and eyelashes is still in the area of interest. Another morphological operation takes place and the limit of eyelids and eyelashes is easier to detect here. Figure 14 shows the new edge map.

The final step concerns total isolation of the iris and determining the pupil parameters (centroid and bounding circle). Efficient localizing of iris is proposed in this part, and the image is ready to get in recognition phase. Figure 15 shows the final isolated iris.

Daugman was suggested his frame work as cropping a 2048 pixel part from any location of the iris. This part will be the main array to perform iris code generation and then matching phase. A sample of cropped rectangle shown in Figure 16

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

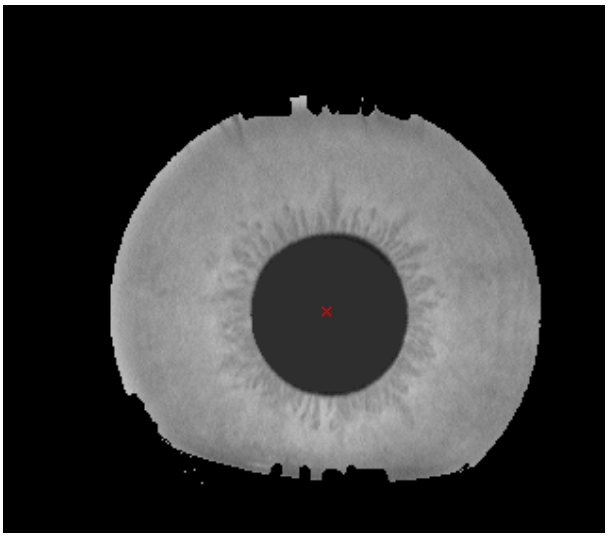


Fig. 15: Final Segmented Iris

Figure 16: Iris Part to be Recognized

IX. CONCLUSION

Iris has the major parameters and features make it important in human identification / verification. All iris recognition applications currently available and all current pass researches depends on patented Daugman's algorithm. This research implements a program to apply the Daugman's equations, and thus perform the iris identification for sample of irises gotten from CASIA data base.

An algorithm for automatic segmentation illustrated and implemented, which localize the iris region in the eye image and isolate from every things around it (i.e. eyelid, eyelash and reflection areas).

Feature extraction of the iris and application of the iris code in the Human Distance (HD) equation, is very reliable for iris recognition for both applications; identification and verification. The HD is the matching metric, which gave a measure of how is two templates related to each other. The statistical dependence test failure of two irises would result a Not-Match.

Finally, the program has been tested in sample irises and gives a full performance in identification.

REFERENCES

[1] J. Daugman. *How iris recognition works*. IEEE Transactions on circuits and systems for video technology, 14(1):21{30, 2004.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

[2] C. Seal, M. Gifford and D. McCartney, *Iris recognition for user validation*. British Telecommunications Engineering Journal 16(7), pp. 113 -117, 1997
[3] J. Daugman. *New methods in iris recognition*. IEEE Trans. Systems, Man, Cyber-netics B, 37(1):1167{1175, 2007.
[4] J. G. Daugman, *High confidence Visual Recognition of Persons by a test of statistical independence*, IEEE Trans. Pattern Anal. Machine Intell., vol. 15,
pp. 1148-1161, 1993
[5] A.K. Jain, A. Ross, and S. Prabhakar. *An Introduction to Biometric Recognition*. Biometrics, 14(1), 2004.
[6] NSTC Subcommittee on Biometrics. *Iris recognition*. http://biometrics.gov/documents/irisrec.pdf/, 2006.
[7] S. Perreault and P. Hebert. *Median Filtering in Constant Time*. IEEE Transactionson Image Processing, 16(9):2389{2394, 2007.
[8] N. Otsu. A *threshold selection method from gray-level histograms*. Automatica,11:285{296, 1975.
[9] T. Min and R. Park. *Eyelid and eyelash detection method in the normalized iris image using the parabolic hough model and otsu's thresholding method*. Pattern Recognition Letters, 30(12):1138 { 1143, 2009.
[10] M. R. Turner, *Texture descrimination by Gabor functions*, Bio. Cybern., vol. 55, pp. 71-82, 1986
[11] Bryan Lipinski. *Iris recognition: Detecting the pupil*. http://cnx.org/content/m12487/1.4/, 2004.
[12] JOHN G. DAUGMAN, *Complete discrete 2-D Gabor transforms by neural networks for imageanalysis and compression*, IEEE Trans. Acoust., Speech, Signal Processing, vol. 36, pp. 1169-179, 1988
[13] H. Davson, *Davson's Physiology of the eye*, 5th ed. London: Macmillan, 1990
[14] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision*. Thomson-Engineering, second edition, 1998.
[15] J. Rohen, *Morphology and pathology of the trabecular meshwork*, in The Structure of the Eye, Smelser, Ed. New York: Academic Press, pp. 335-341, 1961
[16] Samal and P. A. Iyengar, Automatic *recognition and analysis of human faces and facial expressions: A servey*, Pattern Recognit., vol. 25, pp. 65-77, 1992
[17] Teuner and B. J. Hosticka, *Adaptive Gabor transformation for image processing*, IEEE Trans. Signal Processing, in press, 1993

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Design and Implementation of Security Framework for Cognitive Radio Networks Resource Management

Obeten O. Ekabua
Department of Computer Science
North-West University, Mafikeng Campus,
Private Bag X2046, Mmabatho 2735, South Africa

Ifeoma U. Ohaeri
Department of Computer Science
North-West University, Mafikeng Campus,
Private Bag X2046, Mmabatho 2735, South Africa

***Abstract*---Designing and implementing a secure communication for any network is an important issue for the optimal control of resource usage in a resource constrain network environment. Therefore, in this paper, we design and implement a joint authentication and authorization framework by transforming the framework requirement analysis. The framework is a security infrastructure capable of monitoring and controlling access to the limited spectrum resources, dynamically managing data and information in CRN, for a secured communication and quality of service (Q_oS). We explained how the various components in the framework interact to ensure a secured communication and effective access control.**

***Keywords*----Network Management; security; authentication; authorization; access control.**

I. INTRODUCTION

Cognitive radio network is a novel technology designed to alleviate the challenges associated with spectrum shortage. Rapid developments in wireless communication have led to development of Dynamic Spectrum Access (DSA) technology involving licensed and unlicensed users. Secure communication is a salient aspect of any network and has remained unexplored in cognitive radio networks (CRN). Consequently, achieving security in cognitive radio network is thus a huge challenge. The dynamic nature of cognitive radios has introduced weaknesses and vulnerabilities which are capable of affecting the quality of service (QoS) of the network [1,2]. Therefore, the main goal of this research paper is to report on the design and implementation of a joint authentication and authorization framework for CRNs, as a fundamental security infrastructure for access control, and dynamic management of data and information. This security framework can use any form of authentication medium based on network security policy (NSP), either, username, password, pin number and so on. This user profile and security data are supplied to the network management database by registration. Moreover, username and password are used often in this framework design for identification. Often times, users make quick conclusions that, the use of passwords for authentication and authorizations are not reliable and capable of providing a secured communication. When this information is transmitted over the network without encryption, they are prone to attacks because all information and data in the device are exposed. Though, this is not within the context of this research project but however, it is necessary to be mentioned it at this juncture [3].

The design aspect of this paper describes the framework layout and its components using designs and other relevant diagrams for explanations. Authentication and authorization are quite interwoven and often misused. However, the major difference between the two is that authentication deals with the identification of the subject (the client) requesting for connection to the (server), the host connection while authorization determines the access right to the resources (services) available in the network. This makes authentication come first before authorization [4].

II CRN Architecture

Before we introduce the authentication and authorization framework design, it is necessary to first introduce the general design of the CRN

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

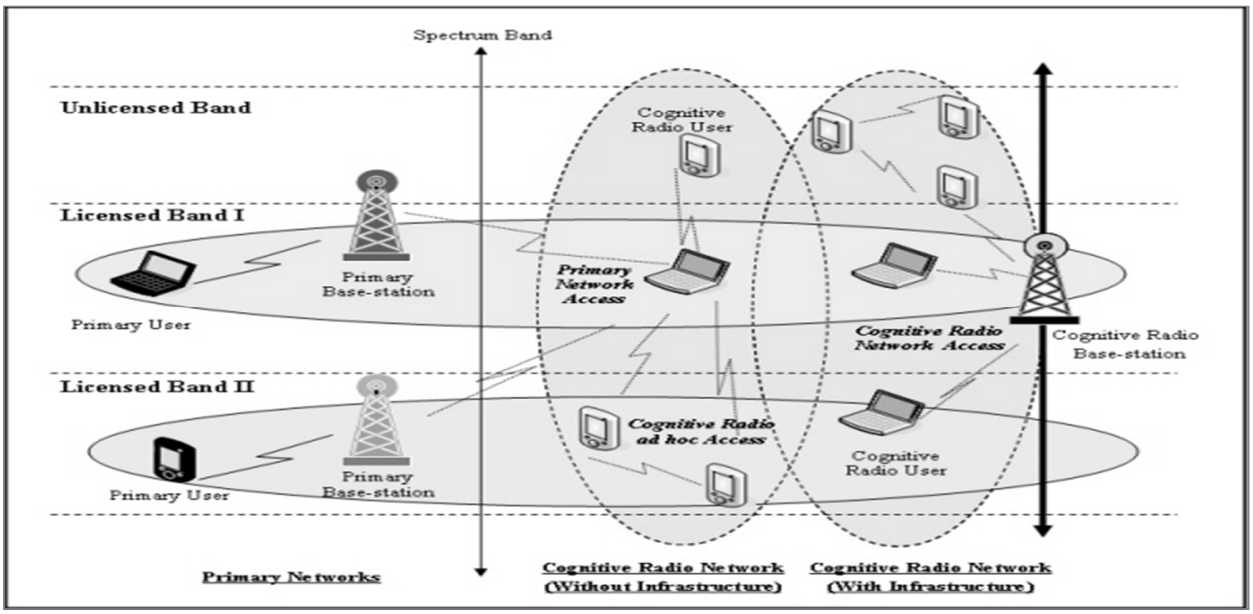


Figure1. Spectrum CRN Architecture and its Interaction

network for a broad view and understanding of the architecture and other relevant components of the concept since this is the architecture (foundation) upon which the research project work is based on. Cognitive radio Network is dynamic and adaptive in nature. The architecture of CRN below shows the different components of, both functional, operational, and hardware, together with the relationship between them. The spectrum band is infinitely renewable, though limited due to its high demand by the secondary users. The Primary user has the legitimate right to a certain spectrum band, whereas, the secondary user do not have the license to operate in a choice band. The primary and unlicensed networks consist of some basic elements which include; primary user, primary base station, cognitive radio user, cognitive radio base station, cognitive radio network access, cognitive radio ad hoc access and primary network access.

However, the Primary user has the license (right) to operate in a specified spectrum band. This access right can only be controlled and monitored by its base-station and unauthorized users are not allowed interfere or affect its operations. Consequently, the Primary base-station is a fixed wireless infrastructure network component that has a spectrum license but do not have any capability for cognitive radio to share the spectrum with other users of cognitive radio. Therefore, the primary base-station may need to have both the primary and cognitive radio protocols to enable primary network access for the cognitive radio users.

Moreover, the spectrum access is allowed for the cognitive radio users only when not occupied by the authorized users because they do not operate with the spectrum license. Therefore, the cognitive radio user capabilities such as; spectrum sensing, spectrum decision, spectrum handoff and cognitive radio MAC, routing and transport protocols are required to enable communication with the base-station and other cognitive radio users as well.

The cognitive radio base-station in Fig. 3 is a fixed wireless infrastructure component that has cognitive radio capabilities and provides single hop connection to cognitive radio users without the license for spectrum access. The cognitive radio users communicate with each other either in a multi hop manner or through a base-station. Consequently, the cognitive radio network architecture in Figure 1 consists of three different types of network access such as: cognitive radio network access, cognitive radio ad hoc access and primary network access with different implementation requirements.

However, in cognitive radio network access, secondary users have the capability to access the cognitive radio base-station in both the licensed and unlicensed spectrum bands. The entire interactions takes place inside the cognitive radio network, therefore access scheme does not depend on the primary network. In cognitive radio ad hoc access cognitive radio users communicate with each other on both licensed and unlicensed spectrum bands via ad hoc connection. They

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

are also capable of building their own access technology through which they can communicate. In primary network access**,** when the primary network is dormant, the cognitive radio users are able to access the primary base-station via the licensed band.

III. CENTRALIZED AND DECENTRALIZED CRNS

This cognitive radio network architecture consists of both the centralized and decentralized cognitive radio network. It shows the position of the primary network and cognitive network in terms of spectrum usage, and communication that exist within the base station [5]. Fig. 2 and Fig. 3 below show the distinction and variation between the two types of cognitive radio network. It indicates the nature of communication existing in the two networks. In a centralized cognitive radio network as shown in Fig. 2, information is disseminated via a service base station which control and manages transfer of messages within the network.

A. Centralized Network Architecture

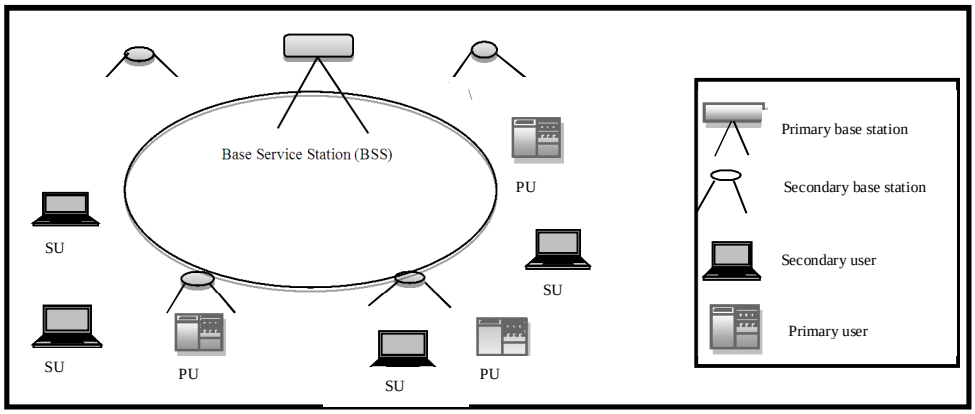


Figure 2. Centralized Network Architecture

B. Decentralized Network Architecture

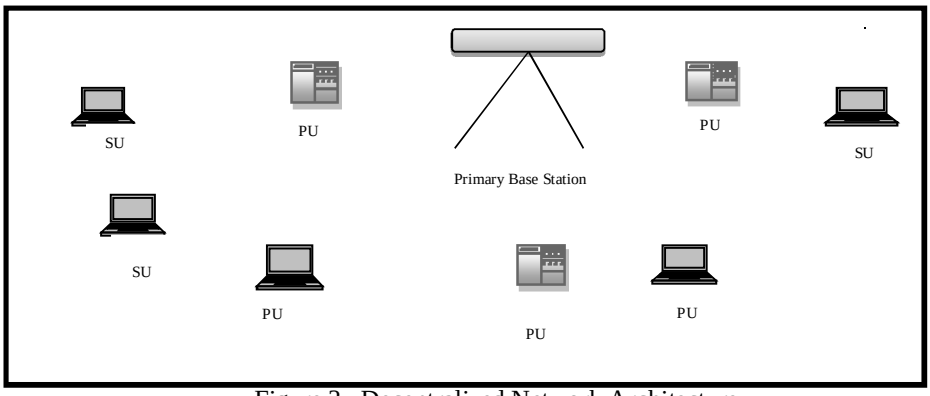


Figure 3. Decentralized Network Architecture

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Data and information are transmitted utilizing radio spectrum frequency bandwidth. Transmission and communication in a decentralized network (fig 3) are transferred directly, but when the devices forming the network are not within a close range, a multi hop is used to enable adequate dissemination of information as used in ad hoc networks.

IV. RATIONAL OF FRAMEWORK

The purpose of the framework in the context of this paper is to ensure a secure communication in cognitive radio network, we use authentication, authorization, as security mechanism, to protect data and information along the line of transmission and also prevent malicious secondary users of the spectrum against network attacks. However, the benefits of are as follows:

1) The framework provides scalability: Typical authentication and authorization configurations depend on a server to or a group of servers to store user name and password. The essence of this is that local databases are not to be built and updated on every router and access server in the network.

2) The framework allows the network administrator configure multiple backup systems. For instance, an access server can be configured to first consult a security server and then the local database before any access is granted.

3) The framework supports standardized security protocols like TACACS +, RADIUS, and Kerberos.

4) The framework provides an architectural capability for configuring two different security measures; authentication, authorization [6].

V. REQUIREMENT ANALYSIS

Requirement analysis firstly specifies the underlying requirement for designing and developing the authentication and authorization framework. The host network is the object, while the client host is referred to as the subject. Authentication concentrates on the subject requesting for connection to the network, while authorization concentrate on the subject requesting for a resource.

When the user dials into an access server which is configured using authentication protocol, the access server and spectrum manager prompts the user to make a user name and password available. The security policy decision point (SPDP) which is the request admission control and handoff point, checks to verify if the user is who he claims to be. The security policy enforcement point (SPEP) ensures that the service management policy is enforced by granting or denying access based on network policy.

The access server verifies a user by requesting for user name and password. This verification process is referred to as authentication. At this point the user may either be denied access or granted access. If authentication is successful then the user can be able to execute commands on the network server. The server then determines the commands and resources that should be made available to the user and specifies the privileges and rights the user should have. This process is referred to as authorization.
However, the framework is developed through four operational stages via: “login”, “connection and resource request”, decision and,” grant” or “deny” access stage.

A. Authentication

Authentication is a security measure in Cognitive Radio Network (CRN) that ensures that entities (users) are truly who they claim to be. This is verified before access to the network is granted. It actually associates a unique identity to each user in CRN, such as user identification name or password as approved by the service security policy. Using these unique forms identification client (users) can freely request for the spectrum resources. It involves the process of verification and validation of users’ identity (ID).

1) Requirement Name: Login
Description: This feature enables communication with the server.
Justification: This feature allows a new window to open for connection request to the server by the client.
2) Requirement Name: Server Request
Description: This request will permit the client access into the network for the service he or she wants to access.
Justification: The framework should request the client identity details by requesting for the user identity (user name and password based on the network configuration, authentication, protocols and security policy enforcement point (SPEP).

3) Requirement Name: Decision
Description: This feature allows the framework to make decision based on the security data and service profile. This stage is handled by the request admission control and handoff which consists of the security policy decision point (SPDP) and SPEP.
Justification: The framework should ensure that the client is who he claims to be, before permission to access the network is granted based on SPEP and SPDP.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

4) Requirement Name: Grant or Deny Access.
Description: The framework should ensures that all the network services and communications are secured from intrusion and unauthorized access.
Justification: The framework should permit all authenticated client to have access to the services available.

B. Authorization

Authorization is a security measure that allows access to only the right entities (users) having the approved privilege to the particular resources requested. Different forms of authorization exist such as; out band authorization, signature authentication and password authentication. Moreover, for any communication (interaction or conversation) involving different parties or entities exchanging information, there should exist, a mutual trust relationship across the multiple domains in CRNS.

1) Requirement Name: Resource Request
Description: This feature will permit the authenticated user, to request for specific services and resources he or she wants to ace ss.
Justification: This framework should validate the users request based on service policies before access is released.

2) Requirement Name: Decision
Description: This feature allows framework to make decision based on the privileges the client has over the resources available in the in the network. This stage is usually handled by the request admission control and handoff domain which consists of SPD and SPEP.
Justification: The framework makes sure that the user (client) has access to only the resources which he or she has the right or privilege to access.

3) Requirement Name: Grant or Deny Access
Description: The framework should ensure that all the network resources are protected from unauthorized users.
Justification: The framework should ensure that all users strictly conform to service policies for authorizations based on the privileges given to the user so as to have access to the services and resources provided by the network.

VI FRAMEWORK DESIGN AND EVALUATION

This research paper presents a detailed design and implementation of a joint authentication and authorization framework by transforming the information from the framework requirement analysis. The framework is a security infrastructure that is capable of monitoring and controlling access to the limited spectrum resources by dynamically managing data and information in CRN, for a secured communication and quality of service (QoS). It is illustrated using components and interface relationships that describe the operation and functionality of the framework. This chapter also explains how the various components in the framework interact to ensure a secured communication and effective access control.

In a decentralized network, mobile devices exist in different locations and communicate in an ad hoc manner with any fixed infrastructure as shown in Figure 3. Data and information are transmitted utilizing radio spectrum frequency bandwidth. Transmission and communication in a decentralized network are transferred directly, but when the devices forming the network are not within a close range, a multi hop is used to enable adequate dissemination of information as used in ad hoc networks.

A. Joint CRN Authentication - Authorization (A-A) Framework

Having designed the authentication and authorization framework separately, it is necessary to also design a joint authentication and authorization (A-A) framework as one security infrastructure or gateway for a CRN. Figure 4 below represents the CRN A-A framework showing the relevant components, and how they interact to form a fundamental security infrastructure for effective dynamic management of data and information in CRN.

Basically, the joint authentication and authorization framework consist of a radio network infrastructure (RNI) and a security policy management center (SPMC). The SPMC In this framework consists of a SPMC agent is installed in each base station to monitor the flow or events within the network. The SPMC agents act like the watch dogs to sense intrusions and malicious attacks. They forwards control messages between the secondary devices and monitor spectrum usage. The SPMC agents are also responsible for service management tasks such as handoff management, secondary user services and all forms of monitoring so that the SPMC is not overloaded.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

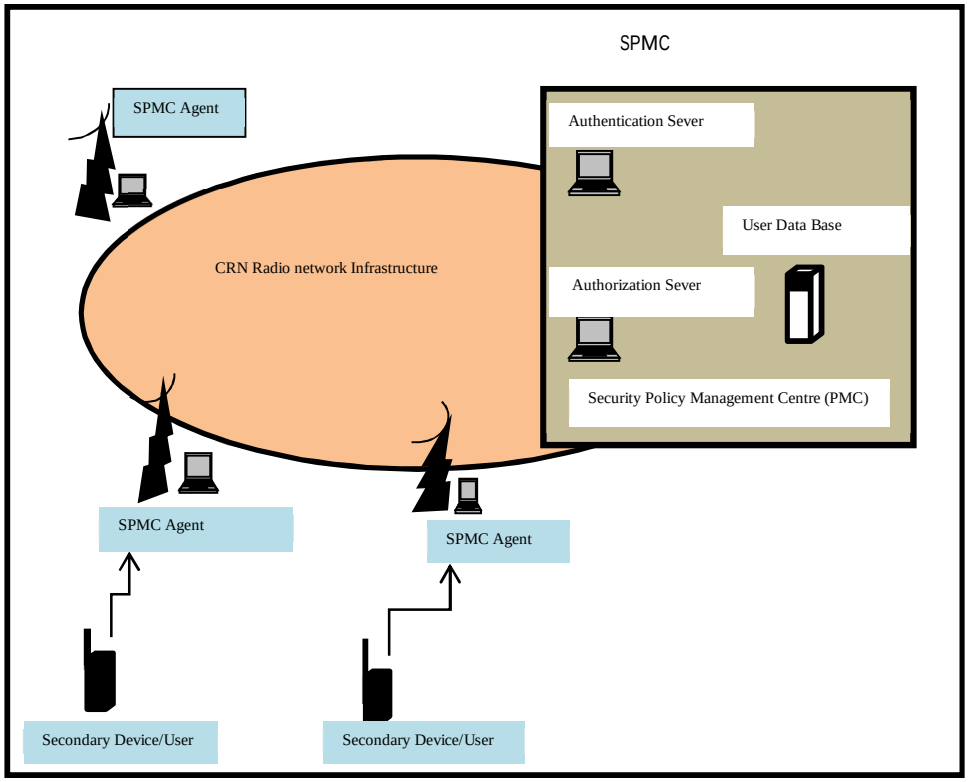


Figure 4. Joint A-A Framework

the parts which includes; an authentication server, (AS), a user database and an authorization server (AS). The authentication server is responsible for authenticating legitimate users. The authorization server is responsible for the spectrum management. Immediately, a user is authenticated and its service requirement is determined to be acceptable, the authorization server authorizes the user by issuing a registration ticket, with which the user can communicate with other users under a close monitoring by the local SPMC agents.

The wireless infrastructure consists of a base station and the mobile switching centers. Moreover,

B. Framework Implication

The reason of this evaluation is to further explain the boarders of the framework. This framework is designed with the assumption that the secondary users or devices adhere to the rules of "inquiring before use" or sensing or listening before use". This means that before the secondary users or devices listen to the control channel allocation information (CAI), notification of the free spectrum channel to utilize before their messages are transmitted for authentication and authorization request.

Software defined radios (SDR) are the key technology behind the CRNs. Therefore, the framework is designed with the understanding that the secondary devices are able to dynamically adjust the radio wave fronts in accordance to the Federal Communication Commission (FCC) spectrum requirement.

Cryptographic methods and public key infrastructure (PKI) required for encryption and decryption are not within the scope of this research project work. We therefore assume that certificate authority (CA) is available to serve the secondary user services such as; issuing public key certificate to the legitimate users of CRNs. Therefore, the verification of public keys and the actual implementation of this framework are among the future work of this research project.

Consequently, for any effort to evaluate this framework, it is necessary to emphasize that this framework is built on the three pillars of secured communication stated below.

1) Privacy

A secured communication or conversation should be private. Only the sender and the receiver (the parties involved) and the devices involved should be able to understand the communication flow.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

Privacy in CRN entails confidentiality and trust relationship. Transmission of data and information among the CR devices in the network must be confidential and the parties or entities involved must be in an agreement of trust to ensure privacy. All security credentials and user registration portfolios to enable access to the available spectrum resources are kept private. In CRN authentication and authorization framework embraces privacy as a major responsibility. It restricts access to message and prevents its contents from being exposed to other users who are not involved in the communication (whether legitimate or malicious users). The aim of privacy standard in the authentication and authorization security framework is to protect the transmission, secure communication and dynamically manage data and information in CRN. This enhances access control and can be achieved by the use of automated encryption.

2) Integrity

A reliable security infrastructure should ensure integrity of the transmitted messages for a secured communication. This ensures that data and information is not altered in an unauthorized manner in transit and that the information received is exactly what is being sent by the transmitter. However, dynamic management of data and information using authentication and authorization security infrastructure ensures that resources are not modified or altered in an unauthorized manner and no third party has unauthorized access to the resources available in the network.

3) Non repudiation

In CRN non repudiation is a feature that establishes the sender of a message or information to the receiver. It works as an accountability measure but also confirms that data and information is authentic and either parties or entities involved in a communication can deny being a part of it. This monitoring and access control feature ensures denial of (resources) data and information to unauthorized users. This is achieved using encryption of a strong access code for user ID which ensures that data and information in CRN are dynamically managed

C. Authentication-Authorization Model

Authentication and Authorization model consists majorly of an engine component called the Authentication and Authorization Engine component. This handles all the decision making activities based on access control policy (authentication and authorization policy).The SPEP for authentication and authorization ensures connection admission control and handoff by enforcing the respective designed policies on the subjects (network users).

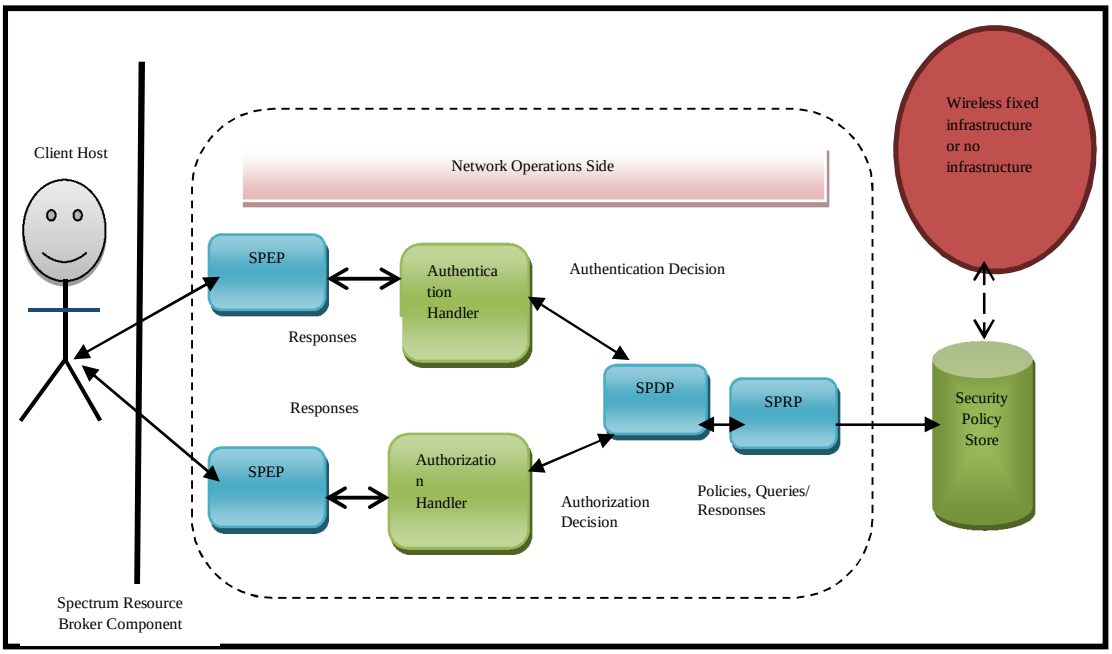


Figure 5. Authentication - Authorization Engine Component

The authentication handler undertakes the decision making process. It decides on who gets connection, for how long and for what purpose. The result from that component is sent to the SPDP for implementation via SPEP based on the stipulated

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

policy, and send confirmation message to the client. The SPRP fetches the policies from the host store. It grants easy access to the policies and helps in selecting the right policy based on request.

D. Spectrum Resource Broker

The Spectrum Resource Broker (SRB) component is the middle man or gateway in the communication line or access path between the client host and the server host and spectrum resources. It manages and controls spectrum resources (data and information). This involves spectrum sharing, spectrum decision and spectrum mobility. All interactions and communications between all cognitive radio networks, both the ones with infrastructure (base stations) and the ones without infrastructure are monitored via the spectrum broker component. It manages all access control mechanisms including; authentication and authorization processes employing the SPEP and SPDP services.
The diagram below indicates what transpires in terms of operations before connections are released from CRNs server and access to spectrum resources is granted.

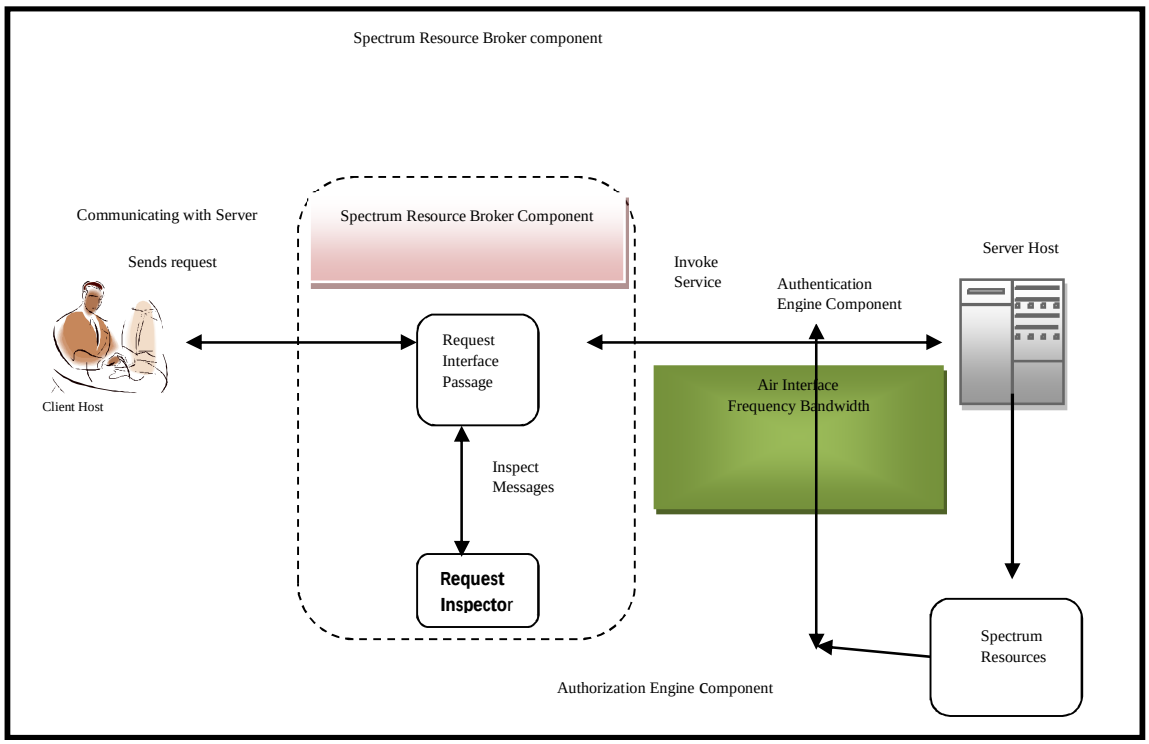


Figure 6. Spectrum Resource Broker Component

E. SRB UML Sequence

The UML diagram describes the sequence of activities in SRB component of CRNs. It shows the operations of its sub components indicating the request and communication (challenge response) AA protocols.

When the client sends a network or resource request it passes through the air frequency bandwidth because of its wireless nature. The request is delivered to the spectrum resource (SRB) broker that consists of the SPEP and SPDP. The SPEP component of the SRB performs the verification activities based on the security service policy (SSP). The message is then validated in line with the SPDP decision and the network service is invoked. The client is given feedback via the SPEP. The access is either granted or denied depending on the verification outcome.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

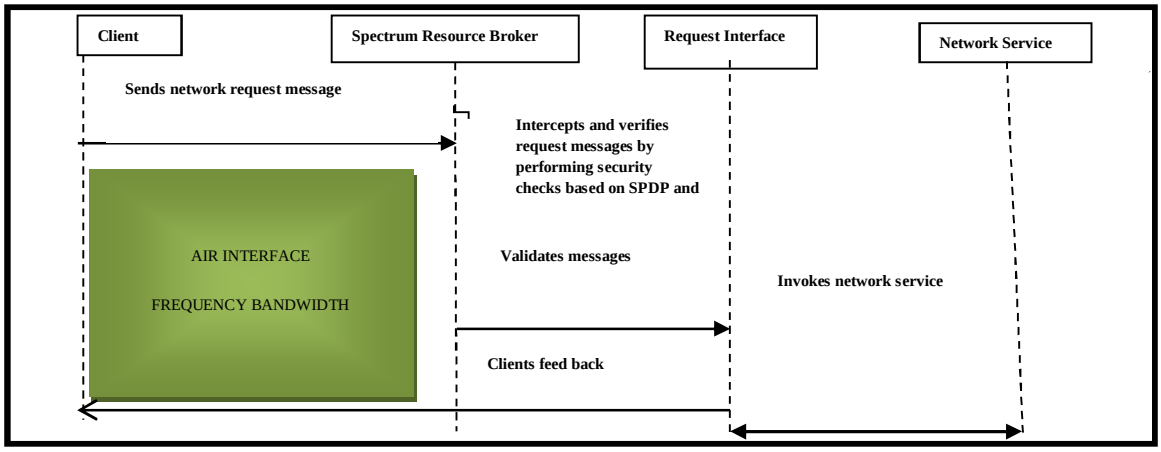


Figure 7. UML Sequence Diagram - Spectrum Resource Broker

F. Security Activity Diagram for A- A Engine Component

The UML Sequence diagram for Authentication-Authorization Engine Component gives a clear description of the relationship and flow of interaction within the A-A engine component and depicts how the service and resource requestor is authenticated and authorized prior to accessing the service and resources and the role each of the components plays in the process of authentication and authorization. Thus, controlling access and dynamically managing data and information in CRN. Authentication takes place before authorization, so it is represented first in the diagram and authorization follows suit.

The Major components of the authentication Engine components such as; the client, the SPEP, the authentication handler (AH), the SPDP, the security policy retrieval point (SPRP) and policy point, are specified in the first column which is the first stage in the sequence.

The arrows pointing downward to the second column specify their corresponding activities and responsibilities respectively, which is the second stage of the diagram. When the client sends the service request message, the SPEP verifies the security details of the client if he is who he claims to be and constructs the authentication decision query and pass over to the SPDP through the authentication handler who certifies the decision query. The SPDP invokes the authentication security policy through the SPRP. The third stage shows the continuous flow of the activities and responsibilities of authentication engine components highlighted in the first column. The arrows pointing to the left hand side in the third column is returning the feedback to the client which is either access granted or denied.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

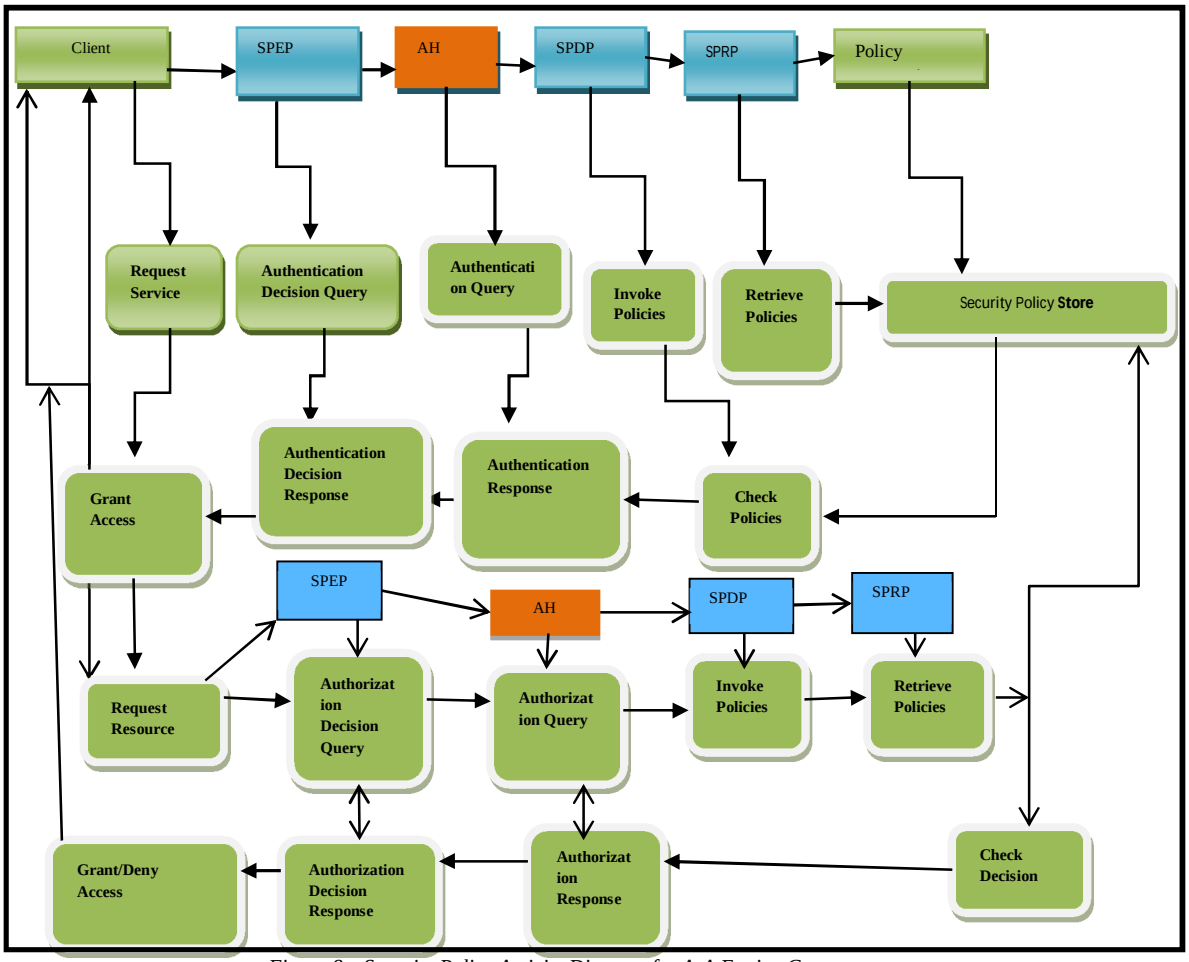


Figure 8. Security Policy Activity Diagram for A-A Engine Component.

The authorization request follows suit in the fourth stage beginning with the resource request from the client which is usually intercepted at the SPEP to perform authorization decisions and passed over to the authorization handler (AH) for authorization query. It then goes over to the SPDP to invoke the security policies which is in turn retrieved from the security policy store by the SPRP. Before the response is returned to the client, the security policy point checks the authorization decision and returns to the authorization handler for response. The decision response is passed over to the client via the SPEP, which is either access granted or access denied.

G. The CRN Usage.

Having understood what CRN and designed its authentication and authorization (A-A) framework, it is also necessary to present the usage diagram for CRN in Fig.12 to show a cross section of the wireless devices in CRN utilizing the spectrum resources. It specifies the several device platforms of CRNS. This means that there is a facility embedded in the devices to enable access to the spectrum resources and enjoy the dividends provided by the network. The service providers are the primary users of the network and they also have end users. The organizations that depend on service providers for the supply and support of the network used to serve their clients constitute the secondary users or end users.

The design clearly explains how the spectrum resources are being utilized and the efficiency of service delivery. Cognitive Radio Network consists of several cognitive radio devices in compatible connection, interacting with each other and the environment to deliver quality services. They interact with the environment in a cognitive cycle which is a core inference mechanism for cognitive devices.

H. Spectrum Management Architecture

The spectrum management architecture is a very important aspect of this research project as it shows the different components that are involved in the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

overall management of the spectrum band. Security as discussed in this research project is an approach for the dynamic management of the spectrum resources (data and information) utilized in CRN. In other words, dynamic management of data and information is majorly about providing a reliable and secured communication of the usage of spectrum resources so as to ensure quality of service (QoS).

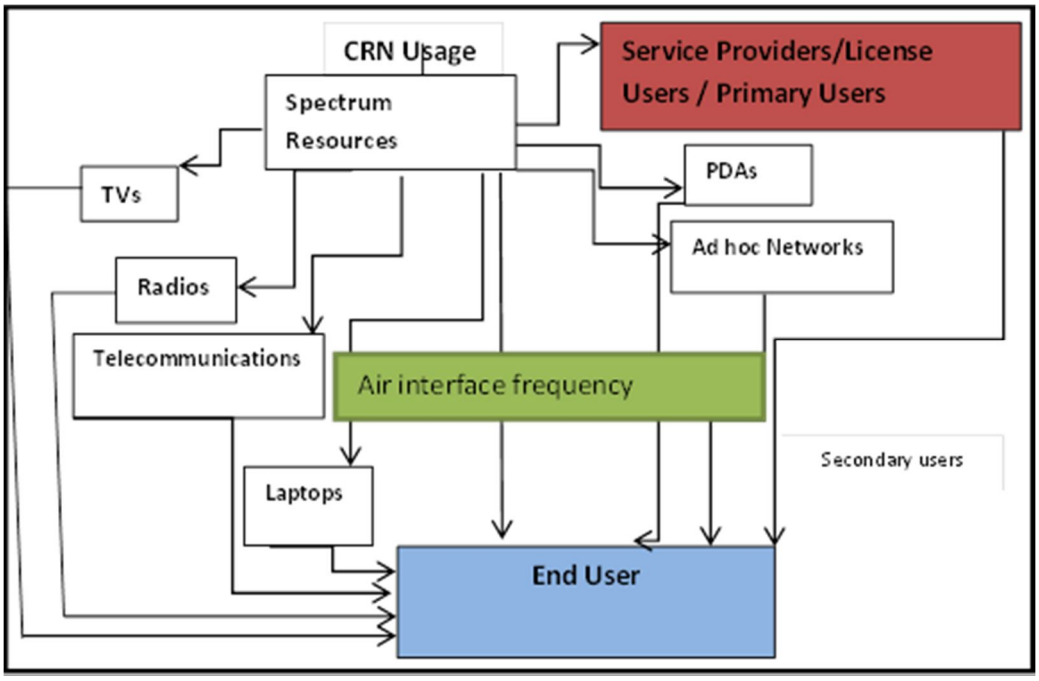


Figure 9. Cognitive Radio Network Usage Diagram

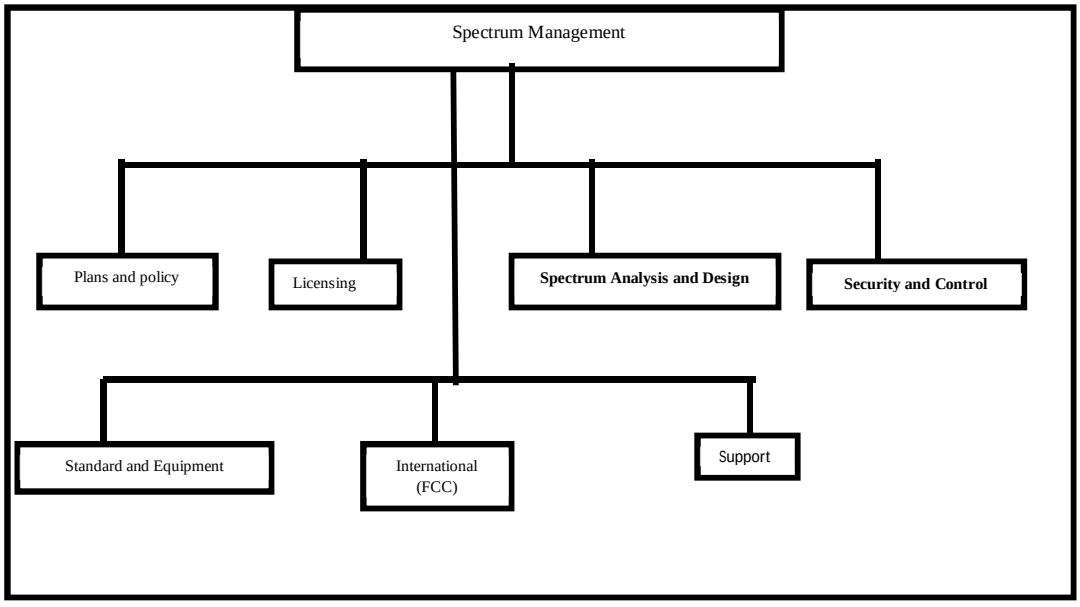


Figure 10. Spectrum Management Architecture

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

The spectrum is similar to become a heterogeneous infrastructure, due to its distributed nature and the high rate of usage and deployment of wireless networks. Therefore, management of data, information and communication in such a distributed environment becomes necessary. The wireless devices operating within both the licensed and the unlicensed spectrum band are controlled and monitored to ensure security. However, the diagram above specifies the relationship and flexibility that exist between the spectrum and CR network employing different components of the spectrum management. The plans and policy entity comprises of, the regulatory policy, spectrum allocation and usage. The licensing entity comprises of, the application using the resource, its terms and condition of registration, review and renewal process. The spectrum analysis entity consists of, the design putting into consideration, interference, avoidance and mitigation. The spectrum control consists of, service policy, enforcement, compliance, control, monitoring and inspection. The standard and equipment identity consists of, authentication, authorization and accounting measures. The international entity consists of, the coordinating body, such as federal communication commission (FCC).

VII. FRAMEWORK IMPLEMENTATION PHASE

The implementation phase demonstrates how CRN clients interact with the system with the aim of proving the concept of authentication and authorization framework for cognitive radio network.

It also shows how access to the services provided by the CR network is controlled and monitored using authentication and authorization access control mechanism as a protective measure against unauthorized and malicious users.

The different interfaces presented in this section indicate the clients' interactions with the system before access is either granted or denied to ensure effective and dynamic management of data and information in cognitive radio network.

A. Jenhosting CRN

The framework is implemented using *Jenhosting* Company (JHC). The company provides numerous services among which are mobile telephony, mobile services, mobile internet and fixed telephony as shown in Fig. 15b. It has numerous clients (subscribers) which include Vodacom, MTN, Celtel, Univen and others. The interface of Fig. 12 shows the CRN home page from which you can navigate to other network domain such as services offered by the network as shown in Fig.16, contact information as shown in Fig.15 and other information about the company as shown in Fig.14, including how to register as shown in Fig.16 and the login outcomes as shown in Fig.18a, Fig. 18b and Fig.18c.

1) Jenhosting CRN Company Home Page

The home page of JENHOSTING Company is the main page of the network, which is the entry point to the Cognitive radio infrastructure. It consists of the login button, the register button, including sites of interest shown in Fig.11 and other vital information about the services rendered by the company.

Figure 11. Cross Section Jenhosting CRN Home page

2) Jenhosting Welcome page

This shows the page that comes up when the new member button is clicked

Figure 12. Jenhosting Welcome Page

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

3) Jenhosting CRN General Information Section

The Fig.13 and Fig.14 interface shows the outcome after the 'About us' and 'Contact us' button has been clicked from the home page. All necessary information about the network operations, services offered, including the contact information is viewed from these domains.

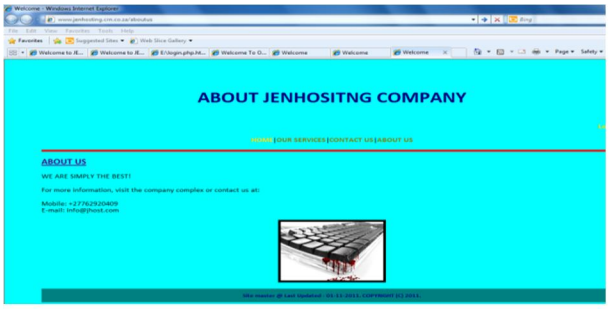


Figure 13. Service Inquiries Page

4) Jenhosting CRN Contact Information Section

This page displays the contact information page when the contact button is clicked.

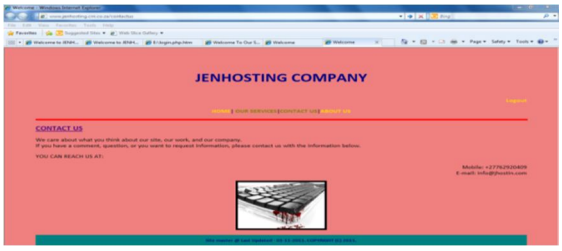


Figure 14a. Contact Page

5) Jenhosting CRN Services

This page displays both the services offered by the cognitive radio network and the available services at the time the service button is clicked.

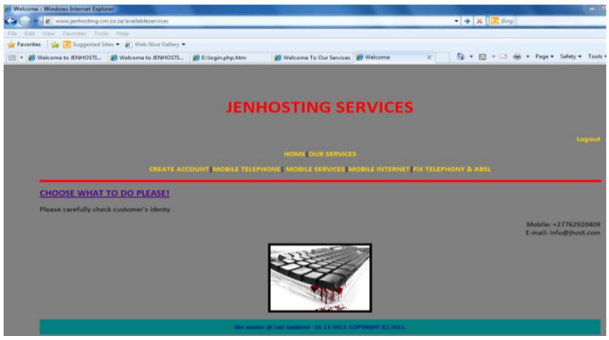


Figure 14b. CRN Services Page

6) Clients e-Registration Section

All the basic information required for the registration of the clients based on the network service policy needed for authentication and authorization are captured from this domain and stored in the data base as shown in Fig. 16.

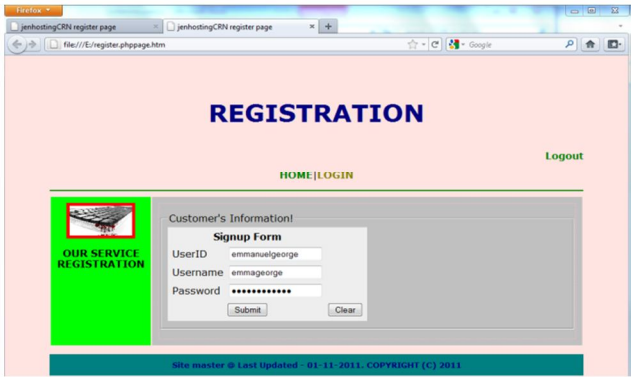


Figure 15. e-Registration Section

7) Jenhosting CRN Database

This represents the authentication and authorization management database and it consists of all the registered clients of the network. The clients name, service name, service ID, password, e-mail and year of registration are clearly specified and stored in this domain for authentication, authorization and security policy services.

Fname	Lname	ServicesName	ServiceID	Physical_Address	Email	Year Reg	Number of Providers	Time
Kehinde	Nkosi	AirTime	SE001	LTT	nb@yahoo.com	2011	2	2011-11-28 11:50:22
Bassey	James	MTN	SE003	Polokwane	baxb@gmail.com	2010	2	2011-11-28 12:05:42
Thembo	Kefi	KC Resturant	SE005	CapeTown	thekefi@aol.com	2010	4	2011-11-28 13:01:50
John	Ebo	Vodacom	SE007	Durban	voda@voda.co.za	2011	3	2011-11-28 15:14:53
Gabriel	Omini	CelTel	SE008	Thohoyandou	cell33@ukcel.co.za	2009	3	2011-11-28 15:16:14
David	Okon	Jaypee Hotels	SE011	LTT	info@jaypee.com	2010	2	2011-11-28 15:17:32

Figure 16. Jenhosting CRN Database.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

8) *Successful Login*

When a request for services is initiated, the client would need to login to the system by supplying identification details (username and password). The details would then be verified and validated from information already stored in the CRN client membership database. A successful login access is granted only if the user is who he claims to be as verified and validated from the database information. In situation where access is not granted, it therefore implies that the request is invalid and an unsuccessful login message would be displayed.

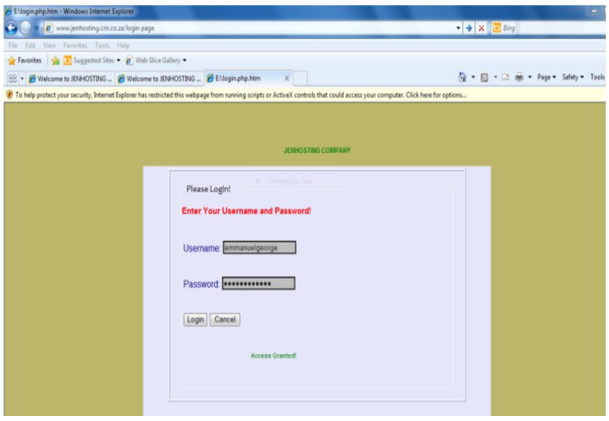


Figure 17a. Successful Login

9) *Unsuccessful Login*

Denial of access to resources during identification of users requesting for services is usually displayed with an unsuccessful login message. This usually happens when a non-registered client is attempting to request for rights of service usage. In such a situation, the system would display unsuccessful login message as a means not to allow malicious intruders into the available services. Unsuccessful login can only be adverted by service requesters registering with the service provider to be allowed access into the CRN resources.

Figure 17b. Unsuccessful Login

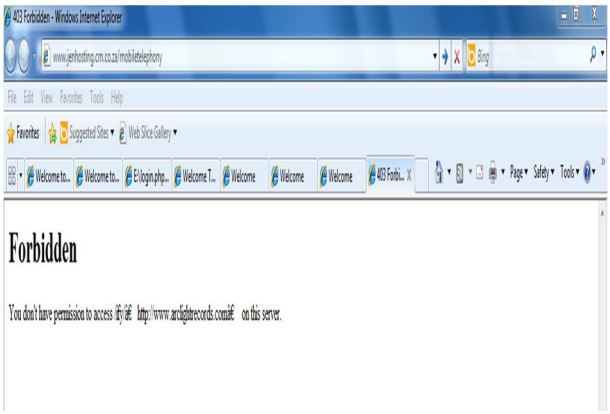


Figure 17c. Unsuccessful Login Section

10) *Delete Account Section*

This implementation phase ensures that no unauthorized user or malicious user masquerades as a legitimate user to gain access to the network server or the resources available in the network for malicious use. This section of the network has the capability to delete the user account and disable the root connections to such users to ensure efficient access control and effective dynamic management of data and information in the specified CR Network.

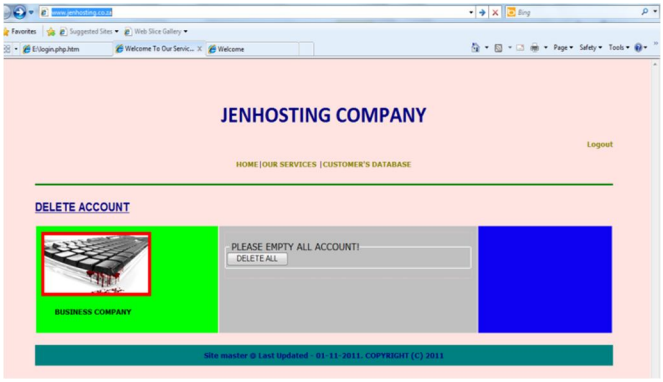


Fig.18: Account Delete Section

B. *Framework Evaluation*

In this paper, we presented an authentication and authorization framework that forms the security infrastructure for access control that can dynamically manage data and information in CRN. It demonstrates how the framework is designed by transforming the artifacts from analysis phase. This paper also has other designs showing the authentication and authorization engine component, the spectrum resource broker component, the UML diagram for authentication and authorization sequence diagram, and the CRN usage diagram. The framework implementation phase consists of various diagrammatic interfaces displaying how the

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

various components of CRN communicate using JENHOSTING cognitive radio network as a model for implementation. Consequently, dynamic management of data and information in CRN provides this reliable security infrastructure as an access control measure to check unauthorized access and all forms of malicious use of the spectrum resources.

Reported in this paper is the design and implementation of authentication and authorization security infrastructure which is able to provide access control and dynamically manage data and information in cognitive radio network to establish control against unauthorized and malicious intruders.

For this controls to be achieved authentication and authorization were introduced. User authentication and authorization is a crucial management component for securing data and information in CRN. Authentication and authorization framework are tightly-coupled mechanisms but also differ in some ways. Authorization process depends on secured authentication mechanism which ensures that a user is who he claims to and thus prevent malicious intruders from gaining access to the secured network resources but also differ in some ways. However, they both offer effective and efficient access control for the dynamic management of data and information in cognitive radio network.

VIII. CONCLUSION

The authentication framework designed in this research report is specifically for cognitive radio networks. The A-A server compares a user's authentication details with the user identification details stored in a database. If the details correspond, the user is granted access to the network. If both information differs the authentication process will fail, then access to the network service is denied.

Authorization is a security mechanism which determines the level of access a specific or particular authenticated user should have to the available and secured network resources. It determines whether a user has the authority to issue certain commands. However, the process enforces policies such as determining what types of activities, resources, or services a user is permitted to perform. The features used are compatible to only the cognitive radio network environment. It is designed to provide efficient and effective dynamic management of data and information in cognitive radio networks. It ensures that data and information are protected to enhance secured conversation.

Summarily, reported in this research is the design and implementation of a security framework that enforces access control policies for optimal spectrum resource management.

REFERENCES

[1]. G. Staple and K. Werbach, “The End of Spectrum Scarcity,” IEEE Spectrum, Vol. 41, No. 3, Mar. 2004, pp.48–52.

[2]. S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," IEEE journal on selected Areas in communications, Vol. 23, No. 2,February 2005.

[3]. Y. Zhou, D. Wu, and S. Nettles. “Architecture of Authentication, Authorization and Accounting for Real Time Secondary Services”, International Jounal of wirwless and Mobile Computing, Vol xx, No x, Jan, 2005.

[4]. O.O. Ekabua, and M.O. Adigun. ”GUISET LogOn: Design and Implementation of GUISET- Driven Authorization Framework,” In Proc. of 1st International Conference on Cloud Computing, GRIDs, and Virtualization, November 21-26, 2010, Lisbon, Portugal pp. 1-6.

[5]. G. Baldini et al. “Security Aspect in Software Defined Radio and Cognitive Radio Networks: A Survey and a Way Ahead,” IEEE Journal, 1553-877x/11/, 2011.

[6]. S. Kumar et al. Ad Hoc Mobile Wireless Networks, www.ubebooks.com-free books and magazines.

Ekabua, Obeten. O. is a Professor and Departmental Chair of the Department of Computer Science in the North West University, Mafikeng Campus, South Africa. He holds BSc (Hons), MSc and PhD degrees in Computer Science in 1995, 3003, and 2009 respectively. He started his lecturing career in 1998 at the University of Calabar, Nigeria. His research interest is in software measurement and maintenance, Cloud and GRID computing, Cognitive Radio Networks, Security Issues and Next Generation Networks.

Ohaeri, Ifeoma U. holds a BSc (Hons) degree in Computer Science in 2006 from the University of Calabar, Nigeria, and another BSc (Hons) degree in Computer Science and Information Systems in 2012 from the University of Venda, South Africa. She is currently pursuing an MSc degree in Computer Science in North West University, Mafikeng Campus, South Africa. Her research interest is in Information Systems and Networks Security, Wireless Networks, and Routing Protocols, Cognitive Radio Networks, and Next Generation Networks.

http://sites.google.com/site/ijcsis/
ISSN 1947-5500

IJCSIS REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway
Assoc. Prof. N. Jaisankar, VIT University, Vellore,Tamilnadu, India
Dr. Amogh Kavimandan, The Mathworks Inc., USA
Dr. Ramasamy Mariappan, Vinayaka Missions University, India
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania
Dr. Junjie Peng, Shanghai University, P. R. China
Dr. Ilhem LENGLIZ, HANA Group - CRISTAL Laboratory, Tunisia
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India
Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India
Mrs Li Fang, Nanyang Technological University, Singapore
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India
Mr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand
Mr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India
Mr. Hayder N. Jasem, University Putra Malaysia, Malaysia
Mr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India
Mr. R. S. Karthik, C. M. S. College of Science and Commerce, India
Mr. P. Vasant, University Technology Petronas, Malaysia
Mr. Wong Kok Seng, Soongsil University, Seoul, South Korea
Mr. Praveen Ranjan Srivastava, BITS PILANI, India
Mr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong
Mr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria
Dr. Riktesh Srivastava, Skyline University, UAE
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia
Mr. Tirthankar Gayen, IIT Kharagpur, India
Ms. Huei-Ru Tseng, National Chiao Tung University, Taiwan

CALL FOR PAPERS
International Journal of Computer Science and Information Security

IJCSIS 2013
ISSN: 1947-5500
http://sites.google.com/site/ijcsis/

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc,), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on
its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embeded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at http://sites.google.com/site/ijcsis/authors-notes .

ISSN 1947 5500
http://sites.google.com/site/ijcsis/